1981

Interpretation of GC-MS data by reverse search and relative retention indices

Christopher Hein
College of William and Mary - Virginia Institute of Marine Science

INTERPRETATION OF GC-MS DATA BY REVERSE SEARCH AND RELATIVE RETENTION INDICES

A Thesis

Presented to

The Faculty of the School of Marine Science

The College of William and Mary in Virginia

In Partial Fulfillment

Of the Requirements for the Degree of

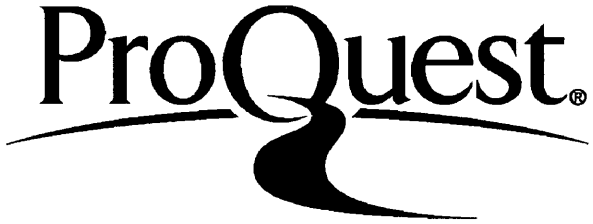
Master of Arts

---

by

Christopher Hein

1981



APPROVAL SHEET

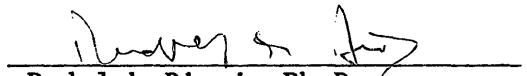
This Thesis is submitted in partial fulfillment of

the requirements for the degree of

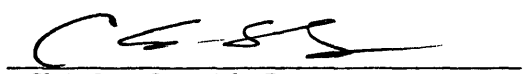
Master of Arts

Christopher Hein

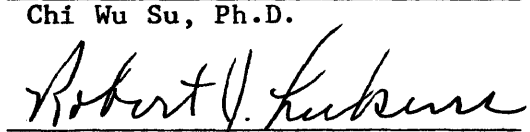
Approved, December 1981



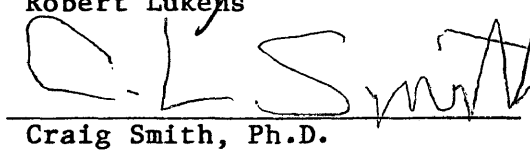
Rudolph Bieri, Ph.D.



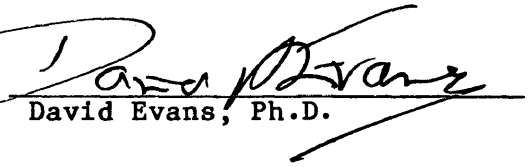
Chi Wu Su, Ph.D.



Robert Lukens



Craig Smith, Ph.D.



David Evans, Ph.D.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I wish to thank Professor R. H. Bieri for his kind guidance and constant support. I am also indebted to Dr. Kent Cueman for providing an introduction into the access of the DuPont GCMS data System, to Buster Blystone and Robert Lukens for generous assistance in programming, to P. C. Gillette of Case Western University for helpful mathematical discussions, and to Molly Lutcavage for assistance in the production of this thesis.

Drafts and final copy of this report were prepared by the VIMS Report Center.

LIST OF FIGURES

ABSTRACT

A technique was developed to interpret GC-MS data from complex environmental samples. Two methods were combined into this technique. The first is a method to evaluate the occurrence of a given reference mass spectrum in an unknown GC-MS data set. The second is a method to optimize the use of gas chromatographic information via relative retention indices. The combination of these two methods provides a thorough, rapid technique to evaluate the potential of occurrence for any number of reference mass spectra and their associated retention indices.

INTERPRETATION OF GC-MS DATA BY REVERSE SEARCH AND RELATIVE RETENTION INDICES

INTRODUCTION

Gas chromatography-mass spectrometry (GC-MS) is a technique used to determine the identity and concentration of components present in complex mixtures of organic compounds. The gas chromatography employed in such analyses may produce several hundred separable peaks. Subsequent mass spectometry provides characterization of the compounds present in the gas chromatographic effluent. GC-MS analysis produces data on the retention time, and mass spectral (mass, intensity) characteristics of each compound present.

GC-MS data must be interpreted to provide meaningful results about the various compounds in a sample mixture. Gas chromatographic and mass spectral characteristics of a given compound are reproducible within limits. Compound identification requires that both GC retention and mass spectral parameters of the unknown compound match those of a standard compound. Identification by GC retention may be limited by the simultaneous or unresolvable elution of two or more compounds. Mass spectral identification is often limited by the similarity of fragmentation patterns in structural isomers. Ideally, GC-MS data analysis is only limited by the simultaneous elution of compounds with similar mass spectra. When this situation occurs, other methods of analysis must be employed.

Factors other than this theoretical limitation in GC-MS data interpretation provide far greater barriers to the complete interpretation of GC-MS data, especially from complex mixtures. As

the number of compounds increases in a sample, the probability for overlapping elution from the GC column also increases. Simultaneous elution produces superimposed mass spectra which in turn increases the difficulty of mass spectral interpretation. Simultaneous elution of various compounds remains a significant problem despite the fact that modern GC-MS analysis incorporates the best available high resolution chromatography. The data produced from an hour long GC-MS analysis may require several days of manual interpretation. Thus, manual interpretation of GC-MS data is an inefficient and time consuming process for routine analyses.

Computer programs have been developed to assist in the interpretation of GC-MS data. The goals of such programs include reducing the time of analysis and increasing the success of identifications. Interpretative programs have in most cases been used on mass spectral information alone. With increasingly complicated applications of GC-MS analysis, efforts have been made to pair retention information with mass spectral data to increase the probability of correctly identifying a compound.

The first computer programs for GC-MS analysis were based on reduced spectra and simulated manual interpretative procedures. Significant features of an unknown spectrum are selected with various criteria. They may then be compared with a library of features selected from reference mass spectra. This strategy of comparing an

unknown spectrum against known characteristics has in historical perspective become known as a 'forward search'.

Work in the field of computerized GC-MS analysis began with two papers co-authored by Crawford and Morrison (1,2). The concept presented in the first paper was to reduce unknown spectra to a small set of most significant mass intensities for comparison against a library of known spectra that had been reduced in similar fashion. The second paper demonstrated that a mass spectrum may be classified according to chemical type via condensed mass spectra.

In their first approach Crawford and Morrison took the six largest peaks of an unknown spectrum and compared this subset with the six largest peaks from 3200 known spectra. The ten library compounds with the greatest similarity to the unknown spectra were reported as possible identifications for the unknown compound. (1) A major problem with this concept is the frequent occurrence of an unknown mass spectrum matching the mass spectra of more than one compound. In this case unpredictable results occur including misidentifications and elmination of the correct spectrum.

By summing fragment intensities of a mass spectrum at intervals of 14 mass units, Crawford and Morrison realized that the resulting condensed mass spectrum was often indicative of a compound's functional group (2). This second strategy supposes that the compounds under investigation as a whole or in part contain saturated aliphatic chains, where 14 mass units signify the loss of CH_2

fragments. Difficulties with this method again relate to the simultaneous presence of an unknown number of spectra, leading to incorrect interpretation.

A number of scientists subsequently developed these concepts and combinations of these concepts for the purpose of identifying unknown spectra. Efficient use of computer time, number of elements versus success in identification, and determining the profile of the average condensed spectra for all functional groups were investigated. (3-7) Two later developments attempted to reduce the frequent problem of superimposed mass spectra. Dromey et al. (8) analyzed mass chromatogram peak shape, and Blaisdell et al. (9) developed a method attempting to separate unknown mixtures of mass spectra into contributing sets.

In 1975, Abramson first pointed out that by reversing the traditional order of library and unknown comparison, the problem of mixed or superimposed spectra could be avoided (10). This strategy had been utilized in part by Pereira et al. (11), and Sweeley et al. (12). In 1976, factor analysis was shown to be applicable to GC-MS data interpretation by Ritter et al. (13). This promising technique is still under development for use in applied GC-MS data analysis.

If a mass spectrum is generated by the simultaneous presence of more than one compound, it represents the simple algebraic sum of all contributing spectral components at their respective concentrations (10). Identification of a known compound in an unknown spectrum would

therefore require that all ions of the known spectrum be present in the unknown mass spectrum.

Pesyna et al. (14) presented a simple two step procedure to evaluate the contribution of a known mass spectrum in an unknown mass spectrum. Intensities of the unknown mass spectrum are divided by the corresponding relative intensities for all masses in the reference mass spectrum, yielding a series of normalized ratios expressed by the variable P below:

$$P_m = \frac{I_m}{RI_m}$$

where I is unknown spectrum intensity, RI is known spectrum relative intensity, and the subscript m indicates the mass, specific for both unknown and reference mass spectrum. The smallest observed P value determines the relative number of times the reference spectrum could be removed from the unknown mass spectrum before a fragment required for the presence of a reference mass spectrum would be exhausted. The intensity of the reference spectrum in the unknown mass spectrum is then obtained by multiplication of the relative number of reference spectra found to fit in the unknown P_{min}, by the sum of RI in each basic reference spectrum. This method assumes that at least one mass in the unknown spectrum is solely from the known compounds, and requires that all elements of the known be present in the unknown mass spectrum. The authors describe a number of additional factors to compensate for distortions in mass spectra from various instruments. The purpose of their study was to test large numbers of known mass

spectra for their possible occurrence in GC-MS data with utmost efficiency.

Gas chromatography (GC) is by itself a powerful analytical method. Separation of a complex mixture by GC leads to the elution of the various components at various points in time. Retention time is subject to variations in chromatographic parameters, thus retention data are usually modified from absolute retention times (injection to elution) and are referred to as relative retention indices. Kovats (15) found that by measuring the locations of eluting peaks in a gas chromatogram relative to a series of n-alkane marker compounds, the effects of variations in absolute retention are minimized. Such retention markers are either coinjected or they may be ubiquitous and clearly identifiable. Any point on the chromatogram that falls within the range of the markers may be expressed via a relative retention index (RRI), adequately defined for linear temperature programming by

$$RRI = \left[\frac{(t_x - t_n) \times 100}{(t_{n+1} - t_n)} \right] + 100n$$

n = number of carbon atoms in n-alkane marker, $(t_x - t_n)$ = retention time difference between unknown and nearest marker with $t_n < t_x$, $(t_{n+1} - t_n)$ = retention time difference of the nearest marker compounds to the unknown.

A relative retention index system will be most accurate if both retention markers and unknown have very similar chromatographic

behavior. Bieri (16) and Lee (17) found it necessary to define an aromatic retention index (ARI) system for work with aromatic compounds. Within one type of compound, the choice of markers is not critical for a retention index system, although a standardized system would facilitate interlaboratory comparison of results. The ARI system described by Bieri used 7 markers (naphthalene, biphenyl, phenanthrene, chrysene, pyrene, perylene and benzo(ghi)perylene) that are given retention indices of 0, 100, 200,...600, respectively.

Recognizing the importance of gas chromatography, Blaisdell and Sweeley (18) developed a GC-MS data interpretive program that incorporated RRI as an identifying parameter. This program observed the characteristic retention of the reference compounds under analysis, and only considered such compounds in their possible retention zones. Each unknown mass spectrum was best accounted for with linear combinations of the reference mass spectra under consideration. The result of this analysis was the printing of concentrations for up to ten library compounds for each unknown spectrum.

Expression of an observed mass spectrum by linear combinations of reference mass spectra takes the form of:

$$MS(observed) = [MS(N) * C(N)] + [MS(M) * C(M)] + ... [MS(un*) C(un)]$$

where $MS(X) = [M(i)*RA(i)] + [M(j)*RA(j)] + [M(k) * RA(k)]...$
(for X = reference mass spectrum N, M,...) and

M(x) is any mass occurrence in the spectrum of compound (X)

RA(x) is a relative abundance associated with M (x),

(for x = mass i,j ....)

C(X) is an absolute intensity factor associated with MS(X)

MS(un) is unexplained mass spectrum

C(un) is intensity of unexplained mass spectrum

Blaisdell and Sweeley's approach was to adjust C(N), C(M), C(L)... in order to minimize C(un). This concept is illustrated below in Figure 1.

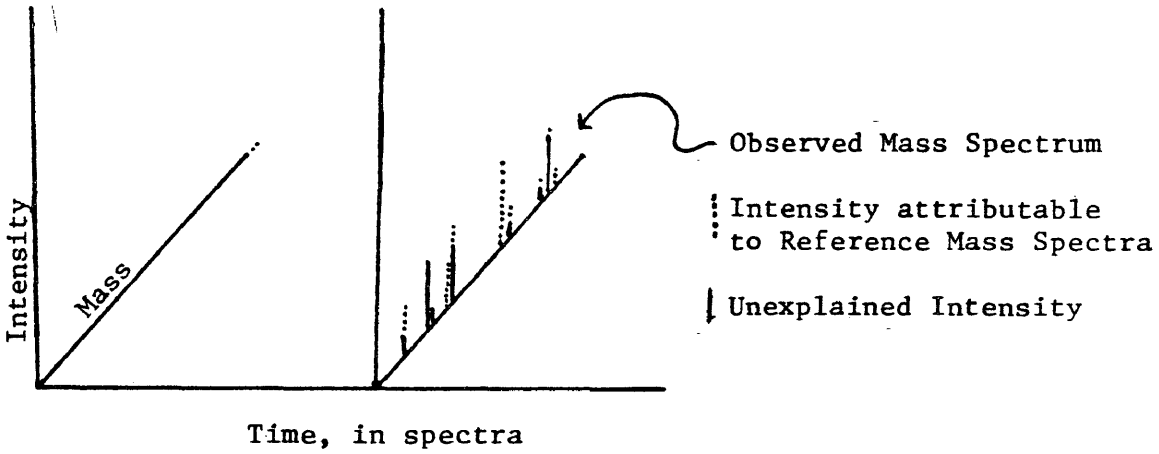


Figure 1. Diagram Illustrating the Strategy of Linear Combination of Reference Mass Spectra to Explain an Observed MS.

Success of this strategy requires prior knowledge of the compounds contributing to the observed mass spectrum. Approximately 30,000 reference mass spectra have been established. Many essentially similar mass spectra occur in similar retention zones. Identification of a compound may be beyond the scope of a given GC-MS analysis under

these circumstances. The linear combination of mass spectra is a method to optimize the intrepretation of an observed mass spectrum with available data. The technique does have a drawback in that it can mask the limitations of a GC-MS analysis.

A compound undergoing gas chromatographic analysis will elute in one concentration vs time peak. In GC-MS analysis, a compound may be located by total ion current, however such measurement is no better than gas chromatographic analysis. Mass spectral analysis in conjunction with gas chromatography allows a greater differentiation between compounds in a complex gas chromatographic separation. Drawing of a compound's fragment intensity vs time allows the designation of time locations where the compound may exist. A common technique to search for the occurrence of a reference spectrum is to select a unique mass for such analysis. A thorough technique to measure the occurrence of a reference mass spectrum was presented by Pesyna et al as previously discussed. This technique determines the area (in scan time *ion intensity) in which a reference mass spectrum may exist in unknown GC-MS data.

The approach of Pesyna et al to GC-MS data interpretation is illustrated in Figure 2. In this manner, a two dimensional plot of the occurrence of a reference mass spectrum may be extracted from three dimensional data.

If more than one peak forms in the allowable intensity for a reference mass spectrum, additional analysis parameters are required.

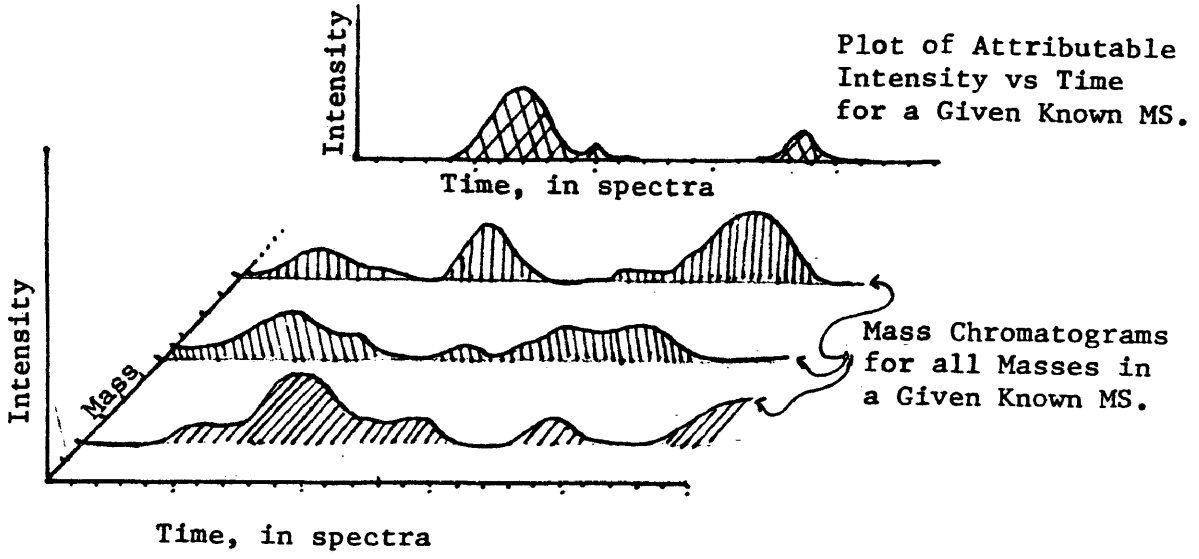


Figure 2. Evaluation of the Occurrence of a Known Mass Spectrum in in a Series of Unknown GC-MS Data via the Method of Pesyna et al. (14).

An additional analysis parameter is available using gas chromatographic retention. A compound may potentially occur within only a limited retention range. If more than one peak may occur in a mass spectrum's intensity vs time plot, within the valid retention region, all GC-MS analysis parameters are satisfied. More definite analysis requires further knowledge of the sample under GC-MS analysis, or another analytical technique.

Pesyna and co-workers' technique would be optimized by integration with a relative retention index system. Such a combination would represent maximized utilization of available data,

and illustrate the limitations of such data. The intent of this study is to implement this combination of reverse mass spectral search with retention indices.

EXPERIMENTAL

Raw GC-MS data were generated with a DuPont 21-094B mass spectrometer and data system. The mass spectrometer scanning speed was modified to one scan cycle per 2.3 seconds, representing the fastest possible scanning speed without calibration difficulties. A Varian 2700 gas chromatograph, modified for capillary operation, was interfaced to the mass spectrometer. The DuPont data system incorporated a Hewlett Packard 2100A computer with a 16K 16 bit word memory.

A series of programs were developed in Fortran IV to provide the DuPont data system with reverse search capabilities. The programs were developed in a DOS III environment, using disk files of library data. A HP 7900A dual 1.5M word disk drive, with removable top disk, was employed with the HP 2100A. Reverse search programs and library files were retained permanently on removable disks, but were copied to the lower disk for analyses. In this manner, various removable disks of GC-MS data could be interfaced with the analysis program. Printouts of analytical results were generated during an analysis on a Tektronix 4012 terminal and 4610 hard copy unit.

RESULTS

An automated interpretative system was developed to supplant manual GC-MS data analysis. A reverse search strategy was determined to be preferable to forward search strategy because of the highly complex mixtures of mass spectra that are encountered in the environmental samples under investigation. The procedure allows an evaluation of the presence of a reference spectrum in unknown mass spectra, in conjunction with a relative retention index system for optimal use of qualitative GC data.

Six individuals systems had to be created before the reverse search could be implemented. These included: 1. a reference mass spectra system, 2. a retention index system, 3. a library for storage and retrieval of reference data, 4. a program to purify reference spectra from GC-MS data (Program THREE) for the collection of reference data, 5. a program to control data in the reference file (Program TWO), and finally, 6. the construction of the desired program itself (called Program ONE). Description of these systems follows.

Reference Mass Spectra

The various fragments of a given compound occur in constant intensity ratios to one another under reproduced mass spectral parameters. It is common practice to express the contributions of different fragments in relative abundance (RA) units. The most abundant fragment (the base peak) was assigned a relative abundance of 100. Other fragments were then expressed in percentage points

relative to the base peak. The reference mass spectrum is thus composed of pairs of information on fragment mass and RA. This data may be used to describe a compound's mass spectrum and thereby identify the occurrence of the compound in GC-MS data. Within the linear operating range of the instrument, the absolute intensity of a mass spectrum is proportional to the amount of compound analysed, thus quantification is possible.

Relative Retention Index

Relative retention index (RRI) systems have been routinely used by this laboratory for manual interpretation. The ARI system described by Bieri (16) and the KRI system described by Kovats were incorporated into the reverse search program to facilitate comparison of data. The version of the reverse search program used with the KRI system required modifying the format of the RRI, dropping the least significant figure and adding a digit in the 10^4 place. The systems are otherwise entirely comparable. In the following, emphasis will be on the ARI system.

Library Structure

In order to perform a reverse search for a series of aromatic compounds, records describing the corresponding reference compounds must be available for use by the computer. This was accomplished by creating a disk file that is read by the computer during analysis. This file was designed to be read in records of 128 integer words, one record for each reference compound. The reference file is ordered by

increasing ARI, thus the first record is of naphthalene (ARI=0) the last record is for benzo(ghi)perylene (ARI=600).

Each record in the reference file has 100 integer words reserved for mass and relative abundance data. Each of the 100 words contains two pieces of data: a fragment mass and the corresponding relative abundance. The fragment mass is limited to a range of 1 to 640 amu, and the relative abundance data must be expressed by numbers between 1 and 100. These data are compacted into a single word to reduce both reference file length, and to reduce the length of time necessary to read the disk. The fragment mass and relative abundance data are compacted into a 16 bit integer word (with range of expression from -32768 to 32767) in the following manner: The relative abundance data are reduced by 1 resulting in a value between Ø and 99. Fragment mass is reduced by 320, yielding a possible value of -319 to 320. This reduced fragment mass is then multiplied by 100 and added to the reduced relative abundance. These efforts produce a compacted number describing mass spectral fragment in the range of -31900 (mass of 1 with an RA of 1) to 32099 (mass of 640, RA of 100).

Decoding of the coded fragment data is accomplished by reversal of the compacting procedure while taking the absolute value of the RA remnant. Word 101 of the reference compound record stores the number of fragments in the reference mass spectrum. Word 102 records the ARI of the reference compound multiplied by 10 to allow one digit past the decimal point in integer number. (In use with the KRI system, the multiplication by 10 must be abandoned in order to allow the

additional descriptor in the 10^4 position of these values). Words 104-123 are used to record up to 40 alphanumeric characters of the compound name. Words 103, and 124-128 are available for parameters possibly useful in the future.

The library disk file was created with the Hewlett Packard Extended File Management Package (EFMP). The file was created on Pack Number (PN) 501, using 1800 sectors of disk area. One file, named ARILI, was set up on PN501 using 1200 of the available 1800 sectors. Thus, the limitation of storage space is 1200 reference records per disk. Access to the reference library may be gained with a Fortran program via HP EFMP EXEC commands, as will be discussed later.

Reference Spectra Acquisition

A computer program (Program THREE) was designed to facilitate compilation of reference data. Although it is possible to enter a reference record manually via Program TWO, it is more convenient to gather reference data from GC-MS records. Program THREE recognizes the phenomenon of mass chromatographic (MC) peaks that form due to the occurrence of a compound in a GC separation, and thus is capable of recognizing which masses may be relevant to the mass spectrum of the compound. Dromey et al. (8) described this techique for spectrum purification. The method is defeated when mass chromatographic peaks are distorted due to common fragment masses of neighboring compounds. Distorted mass chromatographic peak profiles make MC peak time to

compound assignment ambiguous, thereby increasing the difficulty of spectral purification.

Program THREE requests the location of the MS nearest the center in which all ions of a reference compound peak. The computer will then select the masses that form intensity peaks during the 5 mass spectra centered around the designated mass spectrum. For this select group of masses, exact peak center location is stored along with the ion counts (averaged for the three spectra nearest the peak center). The operator is then presented with a table of peak center time versus summed abundance data, and asked to define the time interval for inclusion into the reference spectrum. The mass spectrum that may be included into the designated time location is presented to the operator. If the operator rejects the reference mass spectrum, a new time interval may be designated, or a new set of GC-MS data may be selected for an improved mass spectrum. If the operator accepts the mass spectrum, the computer will request the compound's name and ARI. Finally the new reference spectrum is inserted by the computer into the appropriate position in the reference record file.

The majority of reference spectra in this study were extracted from identified unknowns in archived GC-MS data. A smaller number of reference spectra (approximately 100) were taken from the analyses of standards.

Data Control

A second program (Program TWO) was designed to edit the data on the reference disk file. Program TWO allows the user to display, alter, input, or delete reference data. The display function will plot the reference mass spectrum, name and ARI of a specified compound. This allows checking the record and referencing the library data. The alter function of Program TWO allows changing the ARI or compound name in a reference record. It also allows the addition and deletion of fragments in a reference mass spectrum. The input function of program TWO provides for a manual input of a reference record, thus supplementing Program THREE. The input function may be involved in cases where reference compounds are unavailable, but reference data are published. The delete function allows removal of an unwanted record, and subsequent repacking of the reference file over the vacancy.

Reverse Search Program

The program named ONE was created to perform the desired goal of this project: reverse search of unknown GC-MS data with regard to retention. The program is in the most general terms an interaction point between reference records and unknown GC-MS data with periodic status reports of the interaction sent to the user. This program may be divided into three sections. In the first section, certain variables must be defined before the analysis may be performed. In the second section, a repetitive cycle is carried out once for the

interpretation of each reference compound. The third and final section of Program ONE is the closure of the files accessed during analysis. The three divisions, and their respective steps are outlined in Figure 3.

The central interpretative part of Program ONE includes the following basic steps: In the first step, a reference record of a compound is read from disk to computer memory. The second step involves calculating where the compound could occur in the unknown data and thus define a window (a series of unknown mass spectra) for further investigation. The third step involves evaluation of the mass spectra in the analysis window for intensity attributable to the reference mass spectrum. The fourth step is to interpret these intensities for a peak profile (a requirement of GC analysis). Finally, the reference compound name is printed along with any observed peak data. The computer is then free to return back to the first step for the next compound under investigation, until an end of file mark is reached for either the reference or the GC-MS files.

Reverse Search Analysis Parameters

Early development work on the reverse search program demonstrated that some uncertainty existed in both relative index and reference mass spectra. Slight changes in ARI and reference MS are inevitable and must be accommodated. Parameters were designed into Program ONE to deal with these uncertainties in both reference and unknown GC-MS data.

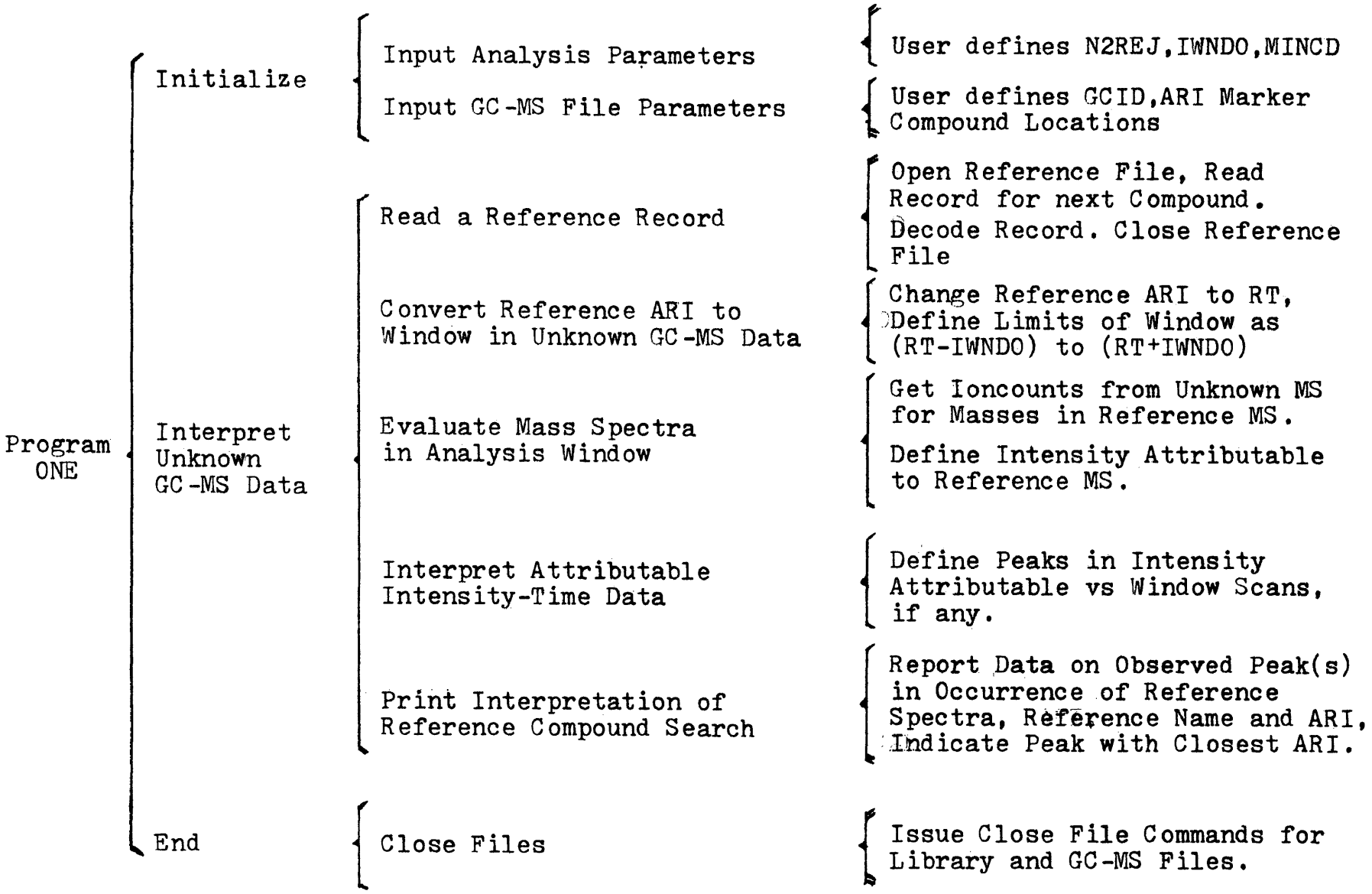


Figure 3. Logical Structure of Program ONE.

Three analysis parameters mut be defined upon initiation of Program ONE: IWNDO, MINCD and N2REJ. Analysis parameters IWNDO and MINCD are factors that allow for uncertainty in ARI and MS data, respectively. The third parameter, N2REJ, was necessary to allow for the presence of minor irrelevant fragments in the reference mass spectra. Values required for this third parameter would decrease to zero if the irrelevant fragments could be recognized and eliminated from the reference data. These three parameters are designed to permit successful analysis using reference and GC-MS data of various qualities. Increasing the value of these parameters allows greater differences to exist in the matching of reference and unknown data.

The variable IWNDO (mnemonic for integer half window) defines the number of spectra to be analyzed on either side of the unknow spectrum of predicted occurrence for the reference compound. Acceptable IWNDO values are from 1 to 12 spectra. Two factors are of importance to the selection of an optimum IWNDO value. First, a wide enough range must be adopted to insure analysis of the compound over the region where it may be significantly represented. Second, the range for analysis must be limited to the region where the compound may possibly occur.

The maximum standard deviation of ARI observed in a 27 compound standard mixture analysed 14 times on one column was 0.63 ARI units (20). Working at a 95% confidence level would necessitate observation of ± 3 standard deviations form the average ARI, or ± 2 ARI units. Two ARI units may translate to a maximum scan time equivalent of 3

spectra, if evaluated within the ARI range of 100 to 200 where each ARI unit corresponds to about 1.4 scans.

The allowable range of IWNDO values is in excess of the statistically required range. This is necessary to allow full base-to-base peak analysis, and to allow flexibility with other systems that may be less accurate than the one used. Experimentation with values of IWNDO has led to common use of 4 spectra for this parameter, thus 9 unknown spectra are evaluated for each reference compound.

Ideally, the fragments produced by the mass spectrometry of a compound will occur in constant intensity relationships with one another. In practice, relative fragment abundances may vary by several percent. In the case of the system used in this experiment, the accuracy with which fragment ions can be measured varies with the square root of the total number of counts (21). Quantification via the technique of Pesyna et al. is based on mass spectral intensities. Error in this quantification in thus proportional to the square root of the absolute intensity of the mass spectral fragments. The variable MINCD was designed into Program ONE to allow a threshold minimum count for the determination of intensity attributable to a reference spectrum contained in an unknown spectrum. Limitation of the quantification to ions with counts of 100 ions or more will provide a relative error of less than 10% with each of the allowed fragments in the unknown data. It is desirable to use as many fragments as possible to define the occurrence of a known spectrum in

unknown GC-MS data, thus insuring a greater resistance to interference from neighboring compounds with masses in common with the reference compound. An optimum MINCD value therefore depends on two factors; first, the error associated with the ion counts used in quantification, and second, the number of intensities that are usable for the quantification. The number of intensities usable for quantification is a value that is determined by the number of fragments in the reference MS, and the number of fragments in the unknown GC-MS data that fall above the MINCD threshold intensity. Experimentally, a MINCD value of 10 has been found to compromise between the error of the single measurement and the accuracy afforded from multi-measurement compound definition.

The variable N2REJ (for the "number at which to reject") was designed to permit successful operation of Program ONE with varying amounts of irrelevant fragments in the reference MS. Irrelevant fragments occur in reference mass spectra due to the possible difficulty in separating relevant from irrelevant fragments. Experience has demonstrated that the majority of irrelevant fragments are of low RA. Fragments of low RA are often invaluable parameters for defining a compound, hence they are included in this analysis. The problem of impurities in principle is simple to solve: a comparison of reference mass spectra from several standard analyses, or from very pure standards, should provide superior reference spectra. This problem was not treated because of the time that would be involved in such multiple standard analyses. The value of N2REJ

defines the sum of missing ions in the reference spectrum (in percent of the total reference mass spectrum) that defines the absence of a reference mass spectrum in an unknown mass spectrum. This places greater importance on the more abundant fragments, i.e. if N2REJ is set at 8%, the absence of four 2% fragments or one 8% (or any other combination totalling N2REJ% or more) terminates the analysis for the mass spectrum of the reference compound in the unknown mass spectrum. N2REJ may be described as controlling qualitative analysis by mass spectrometry, with greater values for this parameter allowing greater dissimiliarity between the reference and unknown MS with a positive identification.

GC-MS File Parameters

Some information about the GC-MS analysis data set is required before a reverse search can be performed by Program ONE. This includes defining the GC-MS data set that is to be analyzed and establishing the retention times for the marker compounds. The GC-MS data files of the DuPont 21-094B data system are accessible via Hewlett Packard Executive Extended File Management Package commands. The GC-MS data files are all recorded on Pack Number 111, with a five character (two letters, three numbers) file name assigned at the initiation of the GC-MS analysis. This file name must be given to Program ONE, along with specific information necessary to locate the marker compounds in the GC-MS data set.

Reference Record Access

Reference data may be read from the reference file, as may GC-MS data, by a Fortran program via the use of Hewlett Packard Executive Extended File Management Package commands, or HP EFMP EXEC. Three such commands are the minimum necessary sequence for such access, and are presented below:

```
CALL EXEC (24,1,IOPNTB,N,ITRBUF,NOTRB,2,IERROR)
CALL EXEC (24,4,IFNA,IPKN,M,ISC,K,IERROR)
CALL EXEC (24,6,IFNA,IRCN,IBUF,IERROR)
```

The first command initiates EFMP abilities by setting up opened file table buffers, track buffers, etc. The second command opens the desired file (IFNA) on the desired pack (IPKN) to allow access to the file by record number in EFMP. The third command creates a copy of the desired record number (IRCN) into a given buffer (IBUF). Complete explanation of EFMP EXEC details are available in a HP DOSIII manual. In practice, checks are made for error messages during EFMP commands. The first two commands listed above need to be executed only once until the file is closed. The effort to gain EFMP access to the files (GCMS and Library) in the programs of this project has been reduced by combining the first two commands into a subroutine called OPEN1 (See Appendix) that only requires the desired file name and pack number be specified.

Once the desired reference record is read into a buffer, it must be decoded for use. This includes unpacking of the mass-relative

abundance data and restoration of the true ARI value, as described in the Library Structure section. A typical reference file record is presented in decoded form in Figure 4.

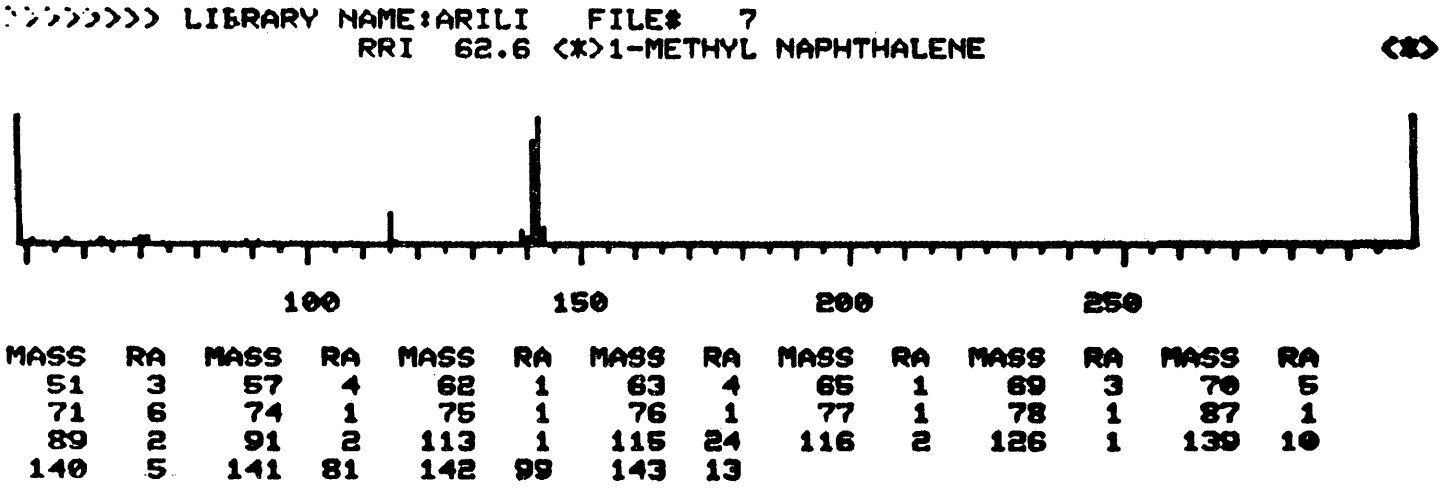
Retention Prediction

The next step in the reverse search analysis is to predict where in the GC-MS data the reference compound could occur. This is performed by converting the characteristic RRI of the reference compound into the equivalent retention time of the GC-MS analysis. A Fortran function CHNGE (See Appendix) was designed to interconvert ARI and retention time (RT). Use of CHNGE requires that the RT and ARI of the retention standards are placed into buffers during initiation of Program ONE for the interconversion calculations. The buffers must be held in common between the main program and CHNGE. The CHNGE function is invoked by the general form:

Y=CHNGE (X,R2S)

whereby, if R2S is positive, X is converted from an RRI value into an RT value and result is stored in Y. If R2S is negative, X is converted from an RT value into a corresponding ARI value and the result is stored in Y. If an X is sent for conversion that falls out of range, a negative CHNGE value is returned, where a -1 specifies before, and -2 specifies after, possible conversion range.

Uncertainty in compound retention and compound peak width requires that a number of mass spectra be investigated before and after the predicted retention for the reference compound. The number



MASS	RA	MASS	RA	MASS	RA	MASS	RA	MASS	RA	MASS	RA	MASS	RA
51	3	57	4	62	1	63	4	65	1	69	3	70	5
71	6	74	1	75	1	76	1	77	1	78	1	87	1
89	2	91	2	113	1	115	24	116	2	126	1	139	10
140	5	141	81	142	99	143	13						

Figure 4. Decoded Reference Record Number 7, from Library ARILI.

of spectra that must be analyzed before and after the predicted retention time is designated by the analysis parameter IWNDO. The window for reverse search analysis is thus defined by the mass spectra between and including (CHNGE (ARI,1)-IWNDO) to (CHNGE(ARI,1)+IWNDO).

Reference Evaluation in Unknown MS

The mass spectra that are included within the window for reverse search analysis must be analyzed one at a time. This is accomplished by a method adapted from the work of Pesyna et al. (14). This method is illustrated with a computer printout of the analysis procedure (Figures 5 to 7) available by using a version of Program ONE called Program HOW, altered to illustrate the intermediate numeric values of the reverse search. Figure 5 shows the results of an analysis of the first spectrum (#78) in the search window for the compound 1-methyl naphthalene.

A negative result was determined in the analysis presented in Figure 5 due to the absence of the reference mass spectrum in the unknown mass spectrum within the specified value of N2REJ. A positive result was later determined in the search for this compound as illustrated in Figure 6. The intermediate step between analysis and user printout is presented in Figure 7.

The search for 1-methyl naphthalene was performed in a relatively simple region of the GC-MS data from analysis of sample STATION J. The reconstructed gas chromatogram of this area of the GC-MS

1-METHYL NAPHTHALENE

REFERENCE SPECTRUM MASS	R.A.	ION COUNT FROM SAMPLE M.S.# 78	P RATIO
143	13	22.	1.7
142	99	80.	.8
141	81	67.	.8
140	5	.	IONCOUNT MISSING, SUM RA
139	10	10.	1.0
126	1	5.	N.A., IONCOUNT < MINCD
116	2	8.	N.A., IONCOUNT < MINCD
115	24	20.	.8
113	1	.	IONCOUNT MISSING, SUM RA
91	2	32.	16.0
89	2	.	IONCOUNT MISSING, SUM RA
87	1	.	IONCOUNT MISSING, SUM RA

ALLOWABLE PERCENTAGE OF TOTAL REFERENCE MASS SPECTRUM (3%) MISSING
REFERENCE SPECTRUM IS THUS DEFINED NOT PRESENT IN THIS MASS SPECTRUM

Figure 5. Search for 1-Methyl Naphthalene in MS #73 of Sample STATION J G3.2

1-METHYL NAPHTHALENE

REFERENCE SPECTRUM MASS	R.A.	ION COUNT FROM SAMPLE M.S.# 84	P RATIO	
143	13	289.	22.2	
142	99	2228.	22.5	
141	81	1878.	23.2	
140	5	117.	23.4	
139	10	239.	23.9	
126	1	22.	22.0	
116	2	49.	24.5	
115	24	429.	17.9	
113	1	20.	20.0	
91	2	35.	17.5	
89	2	32.	16.0	
87	1	15.	15.0	
78	1	.	IONCOUNT MISSING, SUM RA	
77	1	26.	26.0	
76	1	15.	15.0	
75	1	30.	30.0	
74	1	20.	20.0	
71	6	143.	23.8	
70	5	75.	15.0	
69	3	54.	18.0	
65	1	19.	19.0	
63	4	68.	17.0	
62	1	23.	23.0	
57	4	37.	9.2	
51	3	44.	14.7	

SUM OF IONS IN SAMPLE SPECTRUM ATTRIBUTABLE TO REFERENCE COMPOUND	=	SUM OF R.A.s	x	MINIMUM P
2525.		273.		9.2

Figure 6. Search for 1-Methyl Naphthalene in MS #84 of Sample STATION J G3.2

1-METHYL NAPHTHALENE

SPECTRUM	SUM OF IONS IN SAMPLE SPECTRUM ATTRIBUTABLE TO REFERENCE COMPOUND
77	710.
78	.
79	.
80	84.
81	.
82	.
83	1911.
84	2525.
85	262.
86	.
87	.
88	.
89	.

PEAK ANALYSIS

	MAXIMUM	AREA
	80.0	84.
>	83.7	4698.

NOTE: > ASSIGNED TO PEAK NEAREST REFERENCE R.I.

Figure 7. Attributable Intensity within the Search Window for 1-Methyl Naphthalene in Sample STATION J G3.2

analysis is presented in Figure 8-A. A bracket is drawn over the search window of scans 77 to 89.

A second example will be presented from a region of GC-MS data that is more complex. This example will center on the search for fluorene, which is known to occur at an ARI of about 152. This reference retention index is changed into a scan location that the search window is centered on. The search window is then defined as the predicted retention time plus and minus the vaue of the IWNDO parameter established in program initiation. This window is drawn over scans 211-223 in Figure 8-B.

Each unknown mass spectrum in the search window is analysed for the occurrence of the reference mass spectrum. The results of this mass spectral evaluation for two of the unknown mass spectra are presented in Figures 9 and 10. The fluorene search is carried out in a region of high background, as apparent in Figure 8-B. Two major components, and several minor components are obviously present. A number of compounds contribute to the background.

Interpretation of Observations

As the spectra in the analysis window are analyzed, the computer program records the number of ions attributable to the reference spectrum. When all spectra have been analyzed, the sum of ions attributable to the reference compound versus the spectrum number is analyzed (for the occurrence of peaks). If one or more peaks are found, their exact centers are found along with equivalent ARI, peak

```
SEQUEN    21    PAGE    8
GCID  EP   59  STATION J G3.2

#SCANS  150   HRDCPY   NO
%SCALE   70   REZERO   NO
BASE   22251*2**  2
```

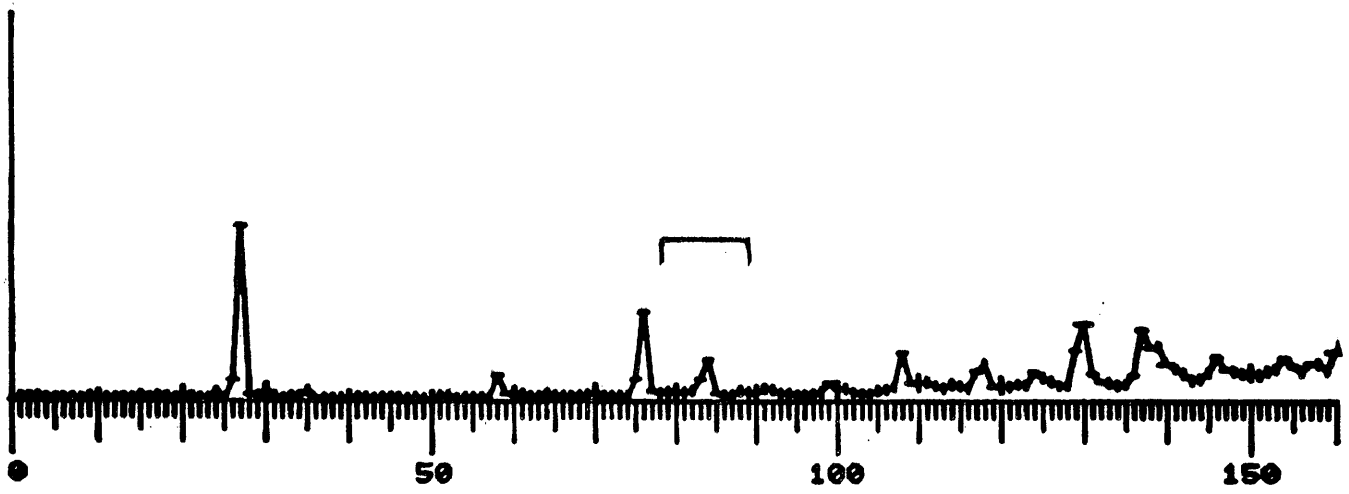


Figure 8A. Reconstructed Gas Chromatogram for Sample STATION J G3.2

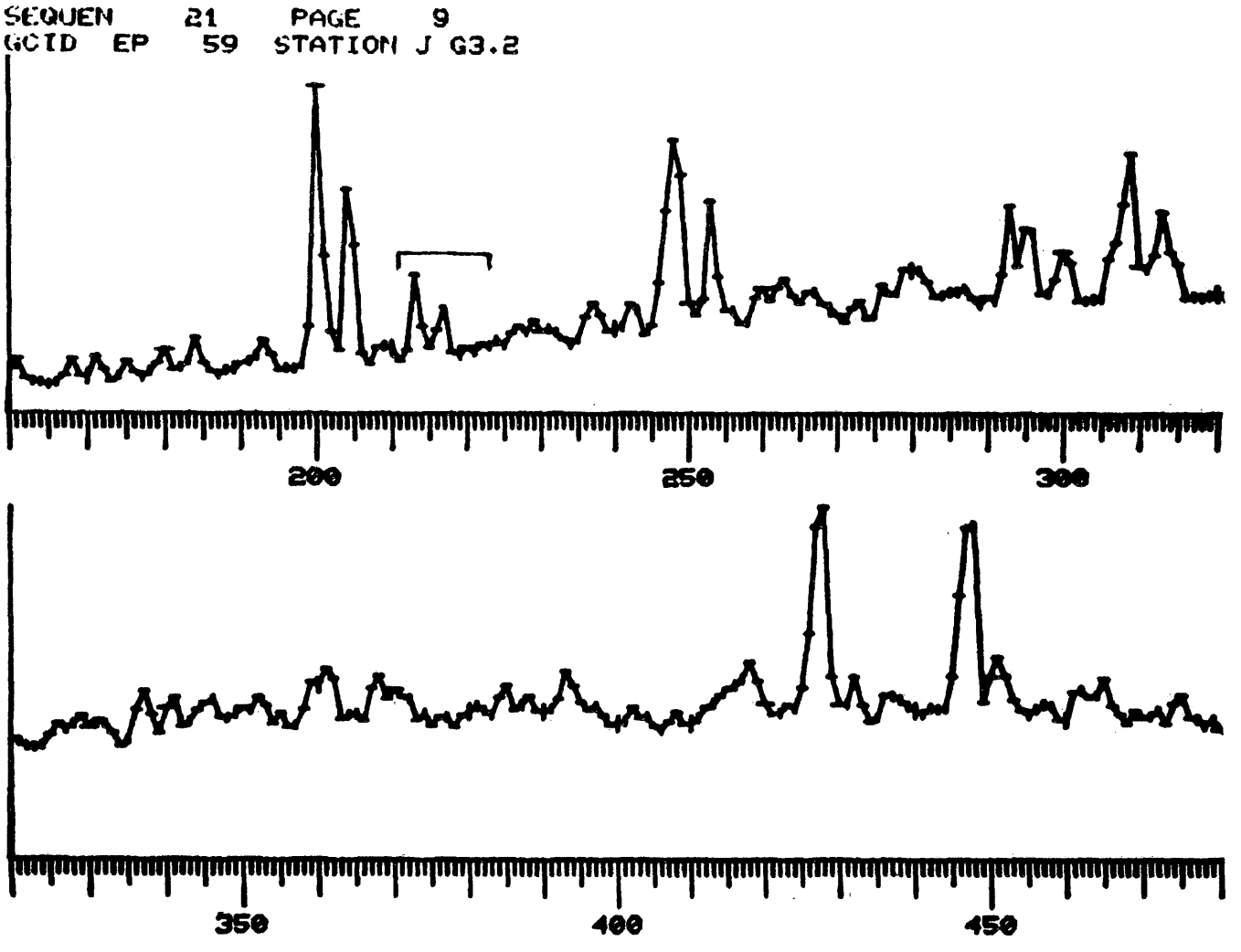


Figure 8B. Reconstructed Gas Chromatogram for Sample STATION J G3.2, Con't.

FLUORENE

REFERENCE SPECTRUM MASS	R.A.	ION COUNT FROM SAMPLE M.S.# 216	P RATIO
168	1	48.	48.0
167	16	265.	16.6
166	100	1246.	12.5
165	88	1154.	13.1
164	10	113.	11.3
163	13	191.	14.7
162	1	23.	23.0
139	6	102.	17.0
115	3	130.	43.3
113	1	40.	40.0
89	1	29.	29.0
87	1	21.	21.0
86	1	18.	18.0
83	8	582.	72.7
82	15	872.	18.1
81	4	304.	76.0
75	1	28.	28.0
74	1	25.	25.0
69	4	428.	107.0
63	3	63.	21.0
62	1	30.	30.0
51	1	49.	49.0

SUM OF IONS IN SAMPLE SPECTRUM ATTRIBUTABLE TO REFERENCE COMPOUND	=	SUM OF R.A.s	x	MINIMUM P
3164.		280.		11.3

Figure 9. Search for Fluorene in MS #216 of Sample STATION J G3.2

FLUORENE

REFERENCE SPECTRUM MASS	R.A.	ION COUNT FROM SAMPLE M.S.# 218	P RATIO	
168	1	12.	12.0	
167	16	85.	5.3	
166	100	664.	6.6	
165	88	547.	6.2	
164	10	72.	7.2	
163	13	113.	8.7	
162	1	30.	30.0	
139	6	63.	10.5	
115	3	93.	31.0	
113	1	20.	20.0	
89	1	38.	38.0	
87	1	9.	N.A.,	IONCOUNT < MINCD
86	1	10.	10.0	
83	8	248.	31.0	
82	15	104.	6.9	
81	4	291.	72.7	
75	1	4.	N.A.,	IONCOUNT < MINCD
74	1	4.	N.A.,	IONCOUNT < MINCD
69	4	338.	84.5	
63	3	29.	9.7	
62	1	7.	N.A.,	IONCOUNT < MINCD
51	1	30.	30.0	

SUM OF IONS IN SAMPLE SPECTRUM ATTRIBUTABLE TO REFERENCE COMPOUND	=	SUM OF R.A.s	x	MINIMUM P
1487.		280.		5.3

Figure 10. Search for Fluorene in MS #218 of Sample STATION J G3.2

boundaries, peak area (in scan time *ion count), and peak with closest ARI to the reference compound. The intermediate step between analysis and user printout is presented in Figure 11.

The intensity attributable to the spectrum of fluorene within its search window is summarized in Figure 11. A graphical output of this data is presented on the reconstructed gas chromatogram in Figure 12. Small peaks were observed at either end of the major fluorene peak. The fluctuation in intensity has three possible causes. First, mass spectral distortions over the elution interval of the compound may have led or contributed to apparent fluctuations in the ideal mass spectrum. Second, the presence of a complex mixture of unresolved compounds (or 'noise') may have led or contributed to the observed results. Third, a compound exhibiting an essentially similar mass spectrum may have eluted in the vicinity of fluorene. This possibility often arises, as illustrated in the next sample for the analysis of 1-methyl phenanthrene.

The reverse search for 1-methyl phenanthrene occurred in a very complex chromatographic region of the GC-MS analysis of sample STATION J. This region is characterized by severe peak overlap, and a large background to peak ratio (approximately 5:1). Manual GC-MS interpretation of regions similar to this require a great deal of skill. The intermediate reverse search results (from Program HOW) are summarized in Figure 13. This data is presented in graphical form in Figure 12.

FLUORENE

SPECTRUM	SUM OF IONS IN SAMPLE SPECTRUM ATTRIBUTABLE TO REFERENCE COMPOUND
211	31.
212	59.
213	56.
214	165.
215	176.
216	3164.
217	6459.
218	1487.
219	283.
220	140.
221	118.
222	182.
223	182.

PEAK ANALYSIS

	MAXIMUM	AREA
	212.4	113.
>	216.9	11961.
	222.5	423.

NOTE: > ASSIGNED TO PEAK NEAREST REFERENCE R.I.

30) FLUORENE	152.1		149.6	212.4	212	813	113
		>	151.9	216.9	213	221	11961
			154.8	222.5	221	224	423

Figure 11. Attributable Intensity within the Search Window for Fluorene in Sample STATION J G3.2

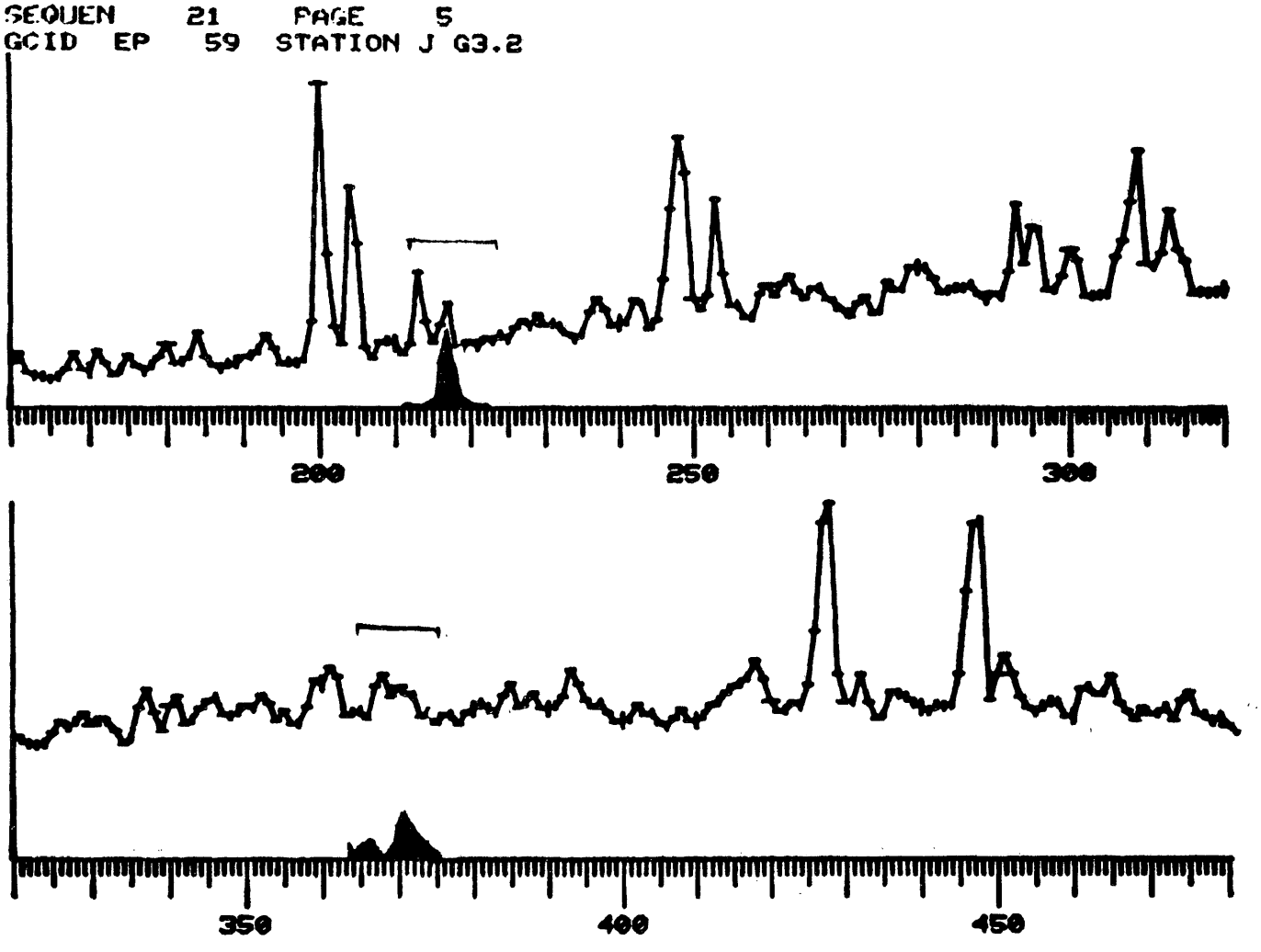


Figure 12. Reconstructed Gas Chromatogram of STATION J G3.2 with Illustrations of Intensity Attributable to Fluorene and 1-Methyl Phenanthrene. Full Scale Represents an Intensity of 62,000 Ions. The Area found Attributable to Fluorene is Represented as a Black Peak Centered around Scan 217. The Area found Attributable to 1-Methyl Phenanthrene is Represented as a Series of Black Peaks Centered around Scan 369.

1-METHYL PHENANTHRENE

SPECTRUM	SUM OF IONS IN SAMPLE SPECTRUM ATTRIBUTABLE TO REFERENCE COMPOUND
363	2172.
364	1227.
365	1825.
366	2095.
367	989.
368	931.
369	3705.
370	4602.
371	3988.
372	3653.
373	1541.
374	843.
375	778.

PEAK ANALYSIS

	MAXIMUM	AREA
	365.7	5988.
>	370.1	19576.

NOTE: > ASSIGNED TO PEAK NEAREST REFERENCE R.I.

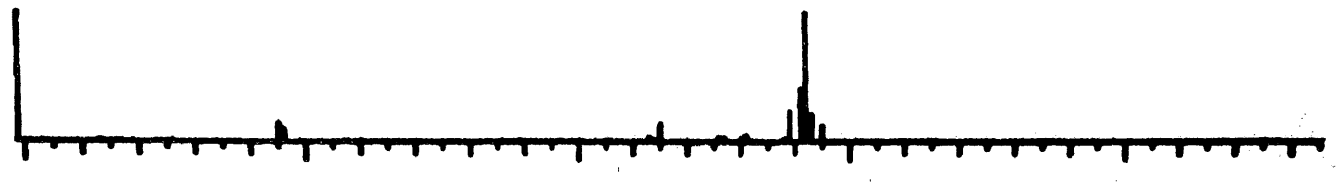
Figure 13. Attributable Intensity within the Search Window for 1-Methyl Phenanthrene in Sample STATION J G3.2

Two major peaks were observed in the search window for the reference spectrum of 1-methyl phenanthrene. This is due to the occurrence of structural isomers of methyl phenanthrene in the analysis window. The structural isomers of a compound often exhibit similar mass spectra and retention. The reference mass spectrum of 1-methyl phenanthrene, and one other compound referred to as me-178 is presented in Figure 14. The exact identity of this me-178 (an abbreviation for a methyl derivative of a compound of M.W. 178) was not established, although it is known to be a methyl phenanthrene or methyl anthracene from a previous GC-MS analysis of a mixture of these compounds. The mass spectra of the two compounds, in addition to the retention index, is very similar. Five isomers of me-178 compounds are presently in the reverse search library. Three more may potentially be defined for methyl phenanthracenes/anthracenes.

Analysis Output

The final step for the reverse search analysis for any compound is to provide a concise output of the computer interpretation. An exaple of the data output for analysis of aromatic compounds is presented in Figures 15A-C. Following the report on each compound, the computer proceeds to the next reference compound until an end of file mark is reached to either the GC-MS data or the reference library. A second example is presented in Figure 16 A-B for the analysis of pesticides and polychlorinated biphenyls in the same GC-MS data set.

>>>>>>>> LIBRARY NAME:ARILI FILE# 69
RRI 240.8 <*>ME-178



MASS	RA	MASS	RA	MASS	RA	MASS	RA	MASS	RA	MASS	RA	MASS	RA
63	2	65	1	75	1	76	2	87	1	88	1	95	15
96	10	110	2	118	1	138	1	139	2	150	1	163	4
164	2	165	14	176	3	177	3	180	3	181	5	188	3
189	24	191	42	192	100	193	22	195	13				

>>>>>>>> LIBRARY NAME:ARILI FILE# 71
RRI 244.1 <*>1-METHYL PHENANTHRENE

MASS	RA	MASS	RA	MASS	RA	MASS	RA	MASS	RA	MASS	RA	MASS	RA
63	1	69	1	75	1	81	2	82	6	93	1	94	9
95	5	96	6	139	1	150	1	151	1	152	1	163	2
164	1	165	7	166	1	176	2	187	2	188	2	189	22
190	9	191	52	192	100	193	14	194	1				

Figure 14. Reference Data for Two Isomers of Structure CH_3-X, where X Represents an Aromatic Compound of Molecular Weight 178.

```
(*) PROGRAM ONE - REVERSE GC-MS SEARCH INCORPORATING RRI.  REVESION 10/21/
81 (*)

-------- ANALYSIS * PARAMETERS --------
SEARCH WINDOW +-  6 SCANS FROM PREDICTED R.I.
REJECT SCAN WHEN  3% REFERENCE SPECTRUM MISSING
DETN CONTRIB VIA   10 OR MORE INTENSE FRAGMENTS

-------- GC-MS FILE PARAMETERS --------
GCID:EP059          STATION J G3.2
 (*) RRI'S)        0      100     200     300     400     500
 (*) @SCAN)     27.0   117.7   308.8   447.5   573.0   705.7   805.8
```

REC#	COMPOUND NAME	LIBRARY DATA ARI		OBSERVED PEAK DATA ARI	SCAN	START@-END@		AREA
1)	NAPHTHALENE (R.S.)	.0	>	-1.0	27.0	25	29	17522
				3.3	30.0	29	31	199
2)	BENZOTHIOPHENE	4.0	>	3.3	30.0	29	31	265
3)	UNKN (SIM 3MECYHEXENON	15.2		8.8	35.0	35	36	81
			>	18.7	44.0	43	45	21
4)	BENZOTHIAZOLE	19.8		-	-			
5)	4-PHENYL-3-BUTEN-2-ONE	36.9		-	-			
6)	2-METHYL NAPHTHALENE	53.8	>	53.8	75.8	74	77	3937
				58.4	80.0	79	81	79
7)	1-METHYL NAPHTHALENE	62.6		58.4	80.0	79	81	84
			>	62.5	83.7	82	86	4698
8)	4-PHENYL-3-BUTEN-2-ONE	89.8	>	89.3	108.0	107	109	62

Figure 15A. Analysis Printout of Program ONE, for Evaluation of STATION J G3.2 with a Reference Library of Aromatic Compounds.

26) 4-PHENYLDECANE	147.5		145.2	204.2	203	207	33413
		>	147.3	208.1	207	211	2016
			150.2	213.7	211	215	3218
27) C3-NAPHTHALENE	147.9	>	147.8	209.0	205	216	5708
28) ME-ACENAPHTHYLENE	150.8		149.6	212.4	211	213	149
		>	151.8	216.8	213	221	9515
29) 3-PHENYLDECANE	151.9		149.9	213.1	211	217	17603
		>	152.5	218.0	217	220	1770
			154.4	221.7	220	223	2281
30) FLUORENE	152.1		149.6	212.4	212	213	113
		>	151.9	216.9	213	221	11961
			154.8	222.5	221	224	423
31) C3-NAPHTHALENE	153.9		152.0	217.0	215	218	623
		>	153.7	220.3	218	223	2158
			155.6	224.0	223	225	138
			156.6	225.8	225	227	97
32) C2-BIPHENYL	158.5		155.6	224.0	224	225	559
			156.7	226.0	225	227	547
		>	158.2	228.9	227	231	3233
			160.0	232.3	231	236	1697
33) 2-PHENYLDECANE	161.3		159.3	231.0	229	233	5508
		>	162.2	236.6	233	239	3677
			163.9	239.8	239	241	873
34) BIPHENYL-AMINE	163.6		162.0	236.1	235	237	1125
		>	163.8	238.4	237	240	2007
			164.7	241.3	240	244	2599

Figure 15B. Analysis Printout of Program ONE, for Evaluation of STATION J G3.2 with a Reference Library of Aromatic Compounds, Con't.

		>	220.6	337.4	335	343	10795
63) ME-DIBENZOTHIOPHENE	222.9		219.7	336.2	335	338	794
		>	223.2	341.0	338	344	5752
			226.5	345.6	344	347	1022
64) 1-PHENYL NAPHTHALENE	227.5	>	226.7	345.8	343	351	3460
			230.9	351.6	350	353	361
65) ME-DIBENZOTHIOPHENE	229.7	>	229.6	349.8	344	352	5892
			232.4	353.8	352	356	1957
66) ME-DIBENZOTHIOPHENE	232.0		229.2	349.4	348	352	5606
		>	232.9	354.5	352	357	3984
			236.1	358.8	357	360	2239
67) ME-178	236.8		234.3	356.3	356	357	370
		>	237.7	361.1	357	364	19759
			240.3	364.8	364	366	2883
68) ME-178	238.3		234.3	356.3	356	357	430
			236.5	359.4	357	360	10130
		>	238.0	361.5	360	364	19228
			241.0	365.6	364	368	7543
69) ME-178	240.8		238.1	361.6	360	364	18801
		>	240.8	365.4	364	368	7159
			244.2	370.2	368	372	16012
70) 4-H-CYCLOPENTA(d,e,f)P	242.4		238.1	361.7	362	364	1469
		>	242.3	367.5	364	371	19483
			245.3	371.7	371	374	3493
71) 1-METHYL PHENANTHRENE	244.1		241.0	365.7	364	368	5988
		>	244.2	370.1	368	376	19576

Figure 15C. Analysis Printout of Program ONE, for Evaluation of STATION J G3.2 with a Reference Library of Aromatic Compounds, Con't.

```
*  PROGRAM ONE - REVERSE GC-MS SEARCH INCORPORATING RRI.  REVESION 10/21/
81 (*)

-------- ANALYSIS * PARAMETERS --------
SEARCH WINDOW +-   6 SCANS FROM PREDICTED R.I.
REJECT SCAN WHEN   3% REFERENCE SPECTRUM MISSING
MIN CONTRIB VIA    10 OR MORE INTENSE FRAGMENTS

-------- GC-MS FILE PARAMETERS --------
GCID:EP059          STATION J G3.2
 (*) RRI'S)      200      300       -1        0        0        0
 (*) @SCAN)    308.8    447.5     -1.0       .0       .0       .0        .0

 /-----LIBRARY DATA---------------\  /---------OBSERVED PEAK DATA-------\
REC# COMPOUND NAME                    ARI      ARI    SCAN  START@-END@      AREA

  24) CL3-BIPHENYL                  203.9    201.2   310.5   309   312        338
                                         >   203.7   313.9   312   316       1023
                                             206.2   317.4   316   319        397
                                             207.9   319.8   319   321        235

  25) DELTA BHC                     213.3    210.0   322.7   322   324        567
                                             211.8   325.2   324   326        550
                                         >   213.2   327.2   326   329        853
                                             215.3   330.0   329   331        552
                                             216.9   332.2   331   334        615

  26) CL3-BIPHENYL                  215.3    211.9   325.3   325   328        466
                                             214.6   329.1   328   330        491
                                         >   216.0   331.0   330   335        841
                                             219.4   335.8   335   337        241

  27) ALPHA ENDOSULFAN              219.5       -       -

  28) TRICHLOROBIPHENYL             227.0  >  227.4   346.7   343   361       3741
```

Figure 16A. Analysis Printout of Program ONE, for Evaluation of STATION J G3.2 with a Reference Library of Pesticides and Polychlorinated Biphenyls.

29) CL3-BIPHENYL	227.1	>	226.8	346.0	343	347	807
			228.2	347.9	347	350	676
			231.0	351.8	350	353	863
30) CL3-BIPHENYL	232.3		228.0	347.7	348	350	386
		>	231.8	352.9	350	355	1827
			234.7	356.9	355	358	615
			235.9	358.6	358	360	265
31) CL3-BIPHENYL	238.5		234.7	356.9	357	358	353
		>	237.7	361.1	358	364	1381
			240.4	364.8	364	366	357
			242.4	367.7	368	369	426
32) HEPTACHLOR	240.9		-	-			
33) CL3-BIPHENYL	243.0		240.4	364.8	364	366	363
		>	242.7	368.0	366	369	499
			244.5	370.5	369	371	362
			245.9	372.5	369	373	743
			246.7	373.5	369	375	1038
34) CL4-BIPHENYL	245.9		-	-			
35) TETRACHLOROBIPHENYL	252.7	>	253.6	383.1	380	385	1801
			256.2	386.8	385	388	738
36) CL4-BIPHENYL	254.4	>	253.7	383.2	388	385	521
			256.4	387.1	385	389	381
37) CL4-BIPHENYL	256.7		253.6	383.2	382	385	564
		>	256.4	387.0	385	389	404
38) CL4-BIPHENYL	258.4		256.4	387.1	385	388	328
		>	257.7	388.8	388	390	226

Figure 16B. Analysis Printout of Program ONE, for Evaluation of STATION J G3.2 with a Reference Library of Pesticides and Polychlorinated Biphenyls, Con't.

DISCUSSION

The presence of an unresolved complex mixture (UCM) in some circumstances defeats the ability of Program ONE to establish the absence of a given reference compound. This is usually due to intensity occurring at all masses during the window of potential occurrence for a given compound. It presently represents a severe limitation with the interpretation of the data. An analogy is the ability to filter a known signal from a noise signal. It is possible to define the maximum potential contribution of the signal to the noise. However, it is impossible to definitively say if the given known signal, or some combination of other signals, contributed to the noise. An unresolved complex mixture is composed of the mass spectra of a finite (but unknown) number of components at various but relatively small concentrations. It is presently impossible to make definitive statements about the composition of the UCM from an unknown sample. Two approaches may be taken in this case. First, chemically separate the components into less complex mixtures, and perform analyses on the fractions. A second approach is to determine the potential intensity attributable to a reference spectrum. If alarming results are produced (i.e. for the concentration of a highly toxic compound), only then would further efforts be expended on a more definitive chemical separation.

A second problem encountered in this reverse search program is the acquisition of high quality reference spectra. Reference spectra quality is a factor that controls the quality of the reverse search

analysis. There are three factors in the quality of a reference spectrum: 1) the inclusion of irrelevant fragments, 2) the exclusion of relevant fragments, and 3) improper assignment of relative abundance relationships for ions of different m/e. Factors 1 and 2 affect the qualitative analysis, and all three factors affect the quantitative analysis performed by Program ONE.

The inclusion of superfluous fragments in a reference mass spectrum will cause the consistent rejection of this compound in the search, except in cases where the same impurity fragments occur along with the reference compound in the unknown data. Such impurity fragments may be due to a compound that co-eluted with the reference compound used for a reference spectrum (in the cases where a pure reference compound was unavailable). Such an impurity may occur intermittently with the referene compound, hence complicating recognition and removal of excess fragments. The analysis variable N2REJ was designed to allow analysis to continue in the event of the above problem. Analysts are free to set the value as low as the quality of the reference mass spectra will permit. If an excessive amount of impurities is present in the reference mass spectrum, the reverse search program will erroneously report the corresponding reference name as not present. A second possible result from a mixed reference spectrum is the reporting of the area where the mixture occurs, with the corresponding (but incorrect) compound name.

Exclusion of relevant fragments from a reference mass spectrum causes a variable degradation in the qualitative, and constant

degradation in the quantitative analysis of the reference compound in the unknown data. Qualitative analysis is affected by such missing fragments when the fragment is a definitive factor in determining the occurrence or absence of the reference compound. This is most likely to be the case if the fragment is in the high mass range, as such fragments tend to be more definitive (22). The evaluation of reference against unknown mass spectra is thus carried out from high to low mass in order to increase analysis efficiency. Quantitative analysis is affected in that the missing fragment abundances will not be included in the area found attributable to the reference compound. This quantitative problem will manifest itself in a consistent percent error in the area attributed to the reference compound.

The improper assignment of relative ion abundance relationships in a reference mass spectrum will also affect the quantification of a reference compound. This factor represents a significant difficulty in producing good quantitative results. In a particular instrument and at a constant partial pressure, the relative ion abundances in a mass spectrum remain constant. In the GC-MS separation, the concentration of a sample is changing as a peak is eluted and this change can distort the observed mass spectrum (23). This problem is illustrated in Figure 17. A reference spectrum may account for only a fraction of the intensity exhibited by the mass spectrometry of the same compound during a GC-MS analysis. Assignment of the true average relationships in a reference mass spectrum will provide the most complete utilization of spectra that are distorted in favor of the

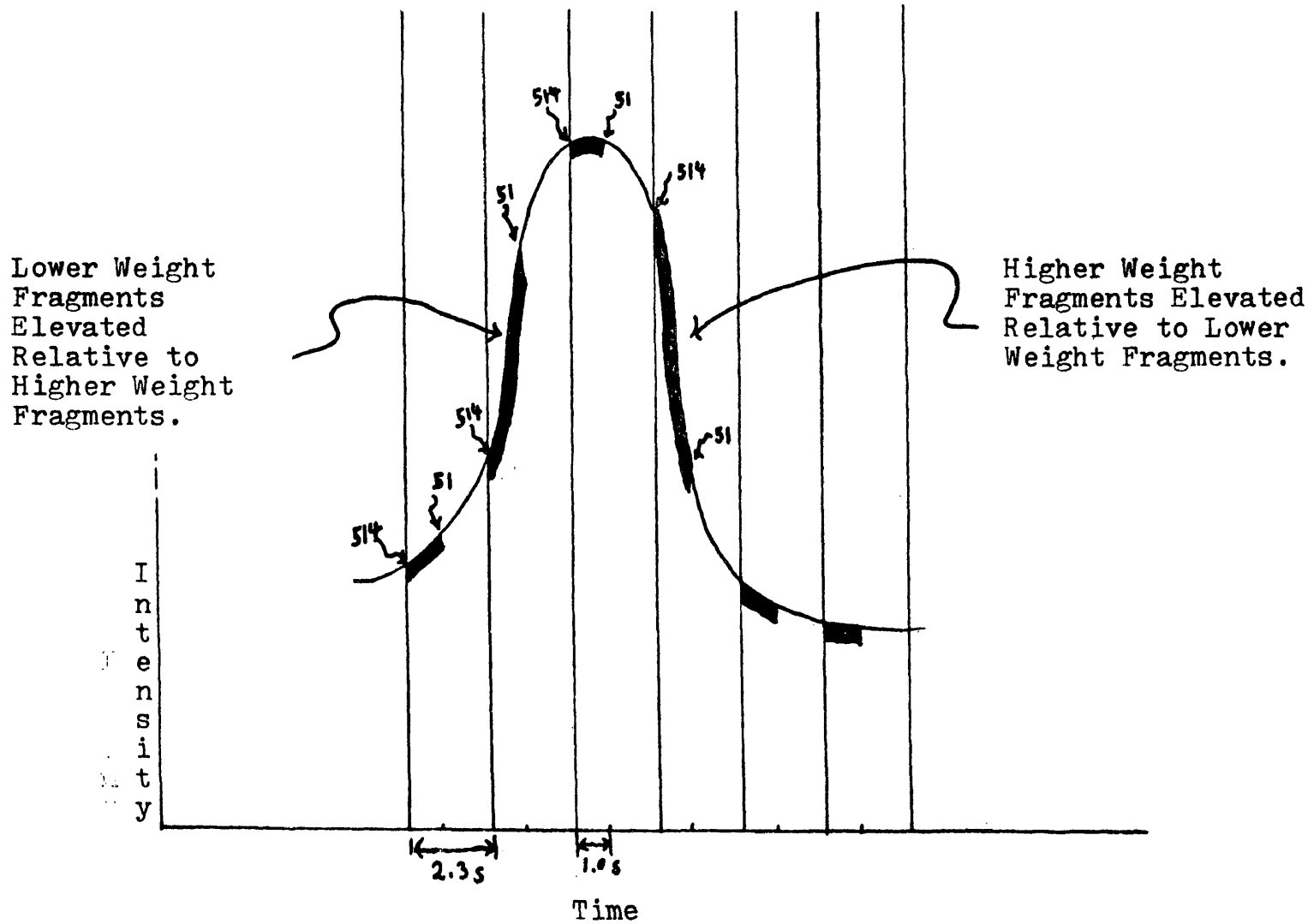


Figure 17. Illustration of the Relationship Between Intensity, Mass and Time in a Gas Chromatographic-Mass Spectrometric Series of Measurements Over the Elution of a Gas Chromatographic Peak. Mass Range on the System used in this Project was 517 to 51 amu. Scan Cycle Time was 2.3 Seconds. Data was Gathered in Each Scan Cycle During a 1.0 Second Interval. The Actual Measurement Period is Indicated by Thick Black Lines.

lower weight fragments during initial buildup of a peak's concentration, and distorted by magnification of the intensities of the higher weight fragments during subsidence of the peak's concentration (assuming higher weight fragments are sampled a finite amount of time before lower fragment masses during collection of the mass spectrum). Improperly assigned relative abundance relationships in a reference mass spectrum may produce unpredictable results, including odd peak shapes of intensity attributable to observed GC-MS data.

It may be possible mathematically to correct the distortions of sampling on an observed unknown mass spectrum. If such a solution is possible, it will be a function of GC peak width, sampling frequency, mass-time relation within a scan, fragment intensity change, and potentially other factors including knowledge of components in the GC peak. Solution of this problem was not attempted, hence the results of Program ONE are described as semiquantitative. Optimization of these semiquantitative results may be possible by creating a response factor for each reference spectrum, plus a certainty factor for the number of observed mass spectra used in quantification and hence express how closely the observed is likely to conform to the averaged reference mass spectrum.

It is possible to optimize the relative abundance relationships in a reference mass spectrum. This was observed by collection of several reference mass spectra from several GC-MS analyses of the same compound. Different quantitative results were provided by these

reference spectra when used to analyze test GC-MS data. The reference mass spectrum that consistently produced the largest quantitative value of the test data was selected as the best representation of the reference mass spectrum. In many cases, two reference spectra produced similar (differing by only a few percent) quantitative results and often neither was consistently more efficient in use of test data. Efforts have been made to develop a routine for optimization of reference mass spectra (24). Such efforts reduce the amount of work required in selecting optimum reference data.

CONCLUSION

Program ONE has demonstrated itself as a useful tool to quickly evaluate the potential occurrence of a series of reference spectra in unknown GC-MS data. The qualitative data thus provided are limited by the quality of the reference data, and in some cases, the complexity of the GC-MS data. The results of Program ONE are frequently too complex to be arrived at manually without extended tedious calculations. The quantitative results are expressed as scan time *ion count and are admittedly only semi-quantitative due to the approximate nature of reference mass spectra.

REFERENCES

1. Crawford, L. R., Morrison, J. D. Anal. Chem., 140, 1464 (1968).

2. Crawford, L. R., and Morrison, J.D. Anal. Chem., 40, 1469 (1968).

3. Knock, B., Smith, I., Wright, D. and Ridley, R. Anal. Chem., 42,1516. (1970).

4. Hertz, H. S., Hites, R. A. and Biemann, K. Anal. Chem., 43,681 (1971).

5. Smith, D. H., Gray N. A. B., Pillinger, C. T., Kimble, B. J. and Eglinton, G. Adv. Org. Geochem., 249 (1971).

6. Heller, S. R. Anal. Chem., 44, 1951 (1972).

7. Smith, D. H. Anal. Chem., 44, 536 (1972).

8. Dromey, R. G. Stefik, M. J., Rindfleisch, T. C. and Duffield, A. M. Anal. Chem., 48,1368 (1976).

9. Blaisdell, B. E. and Sweeley, C. C. Anal. Chim. Acta, 117, 1 (1980).

10. Abramson, F. P. Anal. Chem., 47, 45 (1975).

11. Pereira, W. E., Hoyano, Y., Reynolds, W. E., Summons, R. E. and Duffield, A. M. Anal. Biochem., 55, 236 (1973).

12. Sweeley, C. C., Young, N. C., Holland, J. F. and Gates, S. C. J. Chromatography, 99, 507 (1974).

13. Ritter, G. L., Lowry, S. R., Isenhour, T. L. and Wilkins, C. L. Anal. Chem., 48 591 (1976).

14. Pesyna, G. M., Venkataraghavan, R., Dayringer, H. E. and McLafferty, F. W. Anal. Chem., 48, 1362 (1976).

15. Kovats, Helv. Chim. Acta, 41, 1915 (1958).

16. Bieri, R. Abstracts of 30th Southeastern Regional American Chemical Society Meeting, November 8-10 (1978).

17. Lee, M. L., Vassilaros, D. L., White, C. M., and Novotny, M. Anal. Chem., 51, 768 (1979).

18. Blaisdell, B. E. and Sweeley, C. C. Anal. Chim. Acta, 117, 17 (1980).

19. Tunnicliff, D. D. and Wadsworth, P. A. Anal Chem., 37, 1082 (1965).

20. Personal Communication with C. L. Smith.

21. Faul, H., Ed. 1954. Nuclear Geology, a Symposium on Nuclear Phenomena in the Earth Sciences, John Wiley & Sons, Inc., New York. "Fundamental Considerations, Instruments and Techniques of Detection and Measurement" (Ch. 1).

22. Pesyna, G. M., McLafferty, F. W. Venkataraghavan, R. and Dayringer, H. E. Anal. Chem., 47, 1161 (1975).

23. Junk, G. A. Int. J. Mass Spectrom. Ion. Phys., 8 1 (1972).

24. Blaisdell, B. E. Anal Chem., 49, 180 (1977).

VITA

Christopher Steven Hein

Born in New Haven, Connecticut, June 4, 1957. Graduated from Amity High School in Woodbridge, Connecticut, 1975, B.A. in Chemistry College of Wooster, 1979. M.A. Candidate, College of William and Mary, School of Marine Science, 1979-1981, with a concentration in Chemical Oceanography.

In August 1979, the author entered the Virginia Institute of Marine Science as a research assistant in the Department of Ecology and Pollution.

```
0001   FTN4
0002         PROGRAM ONE
0003   CCCCCC PROGRAM ONE REVERSE SEARCHES GCMS DATA FOR COMPOUNDS
0004   CCCCCC STORED IN LIBRARY CALLED ARILI, IN EFMP FILE PN501.
0005   CCCCCC PN501 READ IN 128 INTEGER WD. RECORDS, WHERE WD 1-100
0006   CCCCCC ARE COMPACTED MASS-REL.ABUND. PAIRS (M=1,650)(R.A.=1,100)
0007   CCCCCC WD 101=NUMBER OF MASS-R.A. PAIRS USED,
0008   CCCCCC WD 102: RRI*10,   WDS 104-123: COMPOUND NAME(ALPHA),
0009   CCCCCC WD 103, WDS 124-128 ARE OPEN FOR FUTURE USES.
0010         COMMON CAT(10),MAT(10),IBUF(128),MAIN(100),CAIN(100)
0011         DIMENSION LIRA(128),TALY(25),IGCID(3),LIRRI(3),NOTE(3)
0012         EXTERNAL EXEC,OPEN,ERASE,CHNGE,GET
0013   CCCCCC RELOCATABLE SUBROUTINES FOR THIS VERSION SAVED AS;
0014   CCCCC OPEN1,PLOT,CHNG5,GET5
0015         DATA IOUT/1/,IN/6/,KSD/5/,KSCAN/1/,LUN/2/,ITRK/190/
0016        &,ISECT/0/,LIRRI(1)/2HAR/,LIRRI(2)/2HIL/,LIRRI(3)/2HI /
0017        &,NOTE(1)/2H >/,NOTE(2)/2H  /
0018         CALL ERASE(1)
0019       1 WRITE(IOUT,111)
0020     111 FORMAT('********* PROGRAM ONE: REVERSE GC-MS DATA SEARCH',
0021        &' USING RETENTION INDICES. *******',//,10X,'DEFINE:',//,
0022        &' IWNDO - # SCANS TO SEARCH FROM PREDICTED OCCURRENCE',/,
0023        &' N2REJ - % REF M.S. ABSENCE FOR COMPOUND'S ABSENCE',/,
0024        &' MINCD - MIN. IONCOUNT TO BE USED IN CONTRIBUTION DETN.',/,
0025        &/,' ** LIMITATIONS **',//,' IWNDO   MIN=1   MAX=12',/,
0026        &' N2REJ   MIN=1   MAX=100',/,' MINCD   MIN=1  NO MAX',//,
0027        &'  NOW INPUT IWNDO,N2REJ,MINCD')
0028         READ(IN,*)IWNDO,N2REJ,MINCD
0029         IF((IWNDO.GT.12).OR.(IWNDO.LT.1).OR.(N2REJ.LT.1).OR.
0030        &(N2REJ.GT.100).OR.(MINCD.LT.1)) GO TO 1
0031       2 WRITE(IOUT,7)
0032       7 FORMAT(' ENTER EXISTING GCID AS FIVE CHARACTERS ')
0033   CCCCCC MAKE SURE DESIRED GCMS ANALYSIS IS ON DISK.
0034         READ(IN,8) IGCID
0035       8 FORMAT(3A2)
```

```
0036          CALL OPEN(111,IGCID,IERROR)
0037          IF(IERROR)2,10,4
0038        4 PAUSE 1
0039       10 WRITE(IOUT,5)
0040   CCCCCC INPUT RRI DATA FOR LOCATING LIBRARY COMPOUNDS.
0041        5 FORMAT('INPUT 1ST STANDARD'S RRI, IE 0, OR 100 ...')
0042          READ(IN,*)MAT(1)
0043          WRITE(IOUT,9)
0044        9 FORMAT(' INPUT SCAN TIMES FOR THE RETENTION STANDARDS,'
0045         &,/,' SMALLEST TO LARGEST.  INPUT -1 AFTER LAST SCAN TIME')
0046          READ(IN,*) CAT(1)
0047          DO 6 I=2,10
0048          READ(IN,*) CAT(I)
0049          IF(CAT(I).EQ.-1) GO TO 20
0050          IF(CAT(I).LE.CAT(I-1)) GO TO 10
0051        6 MAT(I)=MAT(I-1)+100
0052       20 MAT(I)=-1
0053          DO 21 L=(I+1),10
0054          MAT(L)=0
0055       21 CAT(L)=0.0
0056          CALL EXEC(24,6,IGCID,1,IBUF,IERROR)
0057          CALL EXEC(24,5,IGCID,0,IERROR)
0058          IF(IERROR.NE.0)STOP 2
0059          CALL ERASE(1)
0060          WRITE(IOUT,100)IWNDO,N2REJ,MINCD,IGCID,(IBUF(I),I=3,29)
0061         &,(MAT(I),I=1,7),(CAT(I),I=1,7)
0062      100 FORMAT('<*> PROGRAM ONE - REVERSE GC-MS SEARCH',
0063         &' INCORPORATING RRI.  REVESION 10/21/81 <*>',/,
0064         &/,'-------- ANALYSIS * PARAMETERS --------',/,
0065         &'SEARCH WINDOW +-',I3,' SCANS FROM PREDICTED R.I.',/,
0066         &'REJECT SCAN WHEN',I3,'% REFERENCE SPECTRUM MISSING',/,
0067         &'DETN CONTRIB VIA',I5,' OR MORE INTENSE FRAGMENTS',/,
0068         &/,'-------- GC-MS FILE PARAMETERS --------'
0069         &,/,'GCID:',3A2,8X,27A2,
0070         &' <*> RRI'S>',7I8,/,' <*> @SCAN>',7F8.1,2/,
```

```
0071        &" /----LIBRARY DATA---------------\  /---------OBSERVED",
0072        &" PEAK DATA------\",/,'REC# COMPOUND NAME',13X,'ARI     ',
0073        &" ARI    SCAN  START@-END@     AREA',/,)
0074   CCC >>>>>>> 999 IS MAIN LOOP, DO ANALYSIS FOR LIBR-TH COMP. <<<<<<<
0075          DO 999 LIBR=1,1200
0076          CALL OPEN(501,LIRRI,IERROR)
0077   CCCCCC STOP 10 MEANS LIBRARY NOT ON DISK, OR OPEN DAMAGED.
0078          IF(IERROR.NE.0) STOP 10
0079   CCCCCC COPY LIBR-TH LIBRARY RECORD INTO LIRA().
0080          CALL EXEC(24,6,LIRRI,LIBR,LIRA,IERROR)
0081   CCCCCC STOP 77 INDICATES END OF LIBRARY REACHED.
0082          IF(LIRA(102).EQ.-10) STOP 77
0083          IF(IERROR.EQ.0) GO TO 24
0084          WRITE(IOUT,1000)IERROR
0085          PAUSE 3
0086       24 CALL EXEC(24,5,LIRRI,0,IERROR)
0087          ARI=LIRA(102)/10.
0088   CCCCCC CHANGE LIBRARY RRI TO RELEVENT SCAN#.
0089          ICENT=CHNGE(ARI,1)
0090          IF(ICENT+1)30,999,32
0091   CCCCCC ICENT RETURNS -1, BEFORE 1ST RRI, GO TO NEXT COMPOUND.
0092   CCCCCC ICENT RETURNS -2, AFTER LAST RRI, END ANALYSIS (STOP 7)
0093       30 STOP 7
0094   CCCCCC SETUP SCAN RANGE TO SEEK COMPOUND IN.
0095       32 IBSCN=ICENT-IWNDO
0096          IF(IBSCN.LE.0)IBSCN=1
0097          IESCN=ICENT+IWNDO
0098          LASTM=LIRA(101)
0099   CCCCCC STOP 11 MEANS # M/E-R.A. POINTS IN LIBRARY IS>100(BAD).
0100          IF(LASTM.GT.100) STOP 11
0101          NRM=LIRA(101)+1
0102   CCCCCC NRM=NUMBER REF. MASSES +1 FOR BACKWARDS SEARCH LOOP (42)
0103          SUMRA=0.
0104   CCCCCC NOW DECODE LIBRARY MASS-REL.AB PAIRS.
0105          DO 26 I=1,LASTM
```

```
0106          MAIN(I)=320-(LIRA(I)/100)
0107    CCCCCC PAUSE 4 MEANS LIBRARY MASSES NOT IN SEQUENTIAL ORDER.
0108          IF((I.NE.1).AND.(MAIN(I).LE.MAIN(I-1))) PAUSE 4
0109          LIRA(I)=1+IABS(LIRA(I))-IABS((LIRA(I)/100)*100)
0110       26 SUMRA=SUMRA+LIRA(I)
0111          IREJ=(N2REJ*SUMRA)/100
0112    CCCCCC OPEN UP THE GCMS FILE TO READ DATA.
0113          CALL OPEN(111,IGCID,IERROR)
0114          IF(IERROR.NE.0)PAUSE 5
0115    CCCCCC CLEAN OUT COMPOUND CONCENTRATION SCOREBOARD
0116          DO 33 I=1,25
0117       33 TALY(I)=0.
0118    CCCCCC NOW LOOK AT EACH SCAN YOU WANT TO.
0119          DO 36 IWANT=IBSCN,IESCN
0120    CCCCCC ESTABLISH RELATIVE SCAN FOR 'SCOREBOARDS'.
0121          IRSCN=1+IWANT-IBSCN
0122    CCCCCC NZIPS IS #M/E MISSING, IF NZIPS>N2REJ: FORGET SCAN.
0123          NZIPS=0
0124          CALL GET(IGCID,KSD,KSCAN,KFND,IWANT,LASTM,IERROR)
0125          IF(IERROR.EQ.0) GO TO 38
0126          IF(IERROR.EQ.21) GO TO 45
0127          WRITE(IOUT,1000)IERROR
0128          PAUSE 6
0129       38 DIGMIN=10000.
0130          DO 42 J=1,LASTM
0131          I=NRM-J
0132    CCCCCC LOOK AT SPECTRUM STARTING WITH HIGH MASS END, MORE EFFIC?
0133          IF(CAIN(I).EQ.0.) GO TO 34
0134          IF(CAIN(I).LT.MINCD) GO TO 42
0135          DSIG=CAIN(I)/LIRA(I)
0136    CCCCCC DSIG=POTENTIAL QUANTIFYING PARAMETER FOR MASSES>MINCD.
0137          IF(DSIG.LT.DIGMIN)DIGMIN=DSIG
0138    CCCCCC DIGMIN=LIMITING DSIG, USED TO QUANTIFY KNOWN IN SPECTRUM.
0139          GO TO 42
0140       34 NZIPS=NZIPS+LIRA(I)
```

```
0141  CCCCCC IF TOLERANCE OF MISSING IONS IS EXCEEDED, FORGET SCAN.
0142        IF(NZIPS.GT.IREJ) GO TO 41
0143     42 CONTINUE
0144  CCCCCC ST CONC OF LIBR COMPOUND IN TALY().
0145        TALY(IRSCN)=SUMRA*DIGMIN
0146  CCCCCC IF FIRST SCAN OF AN ANALYSIS, SAVE ITS ADDRESS.
0147     41 IF(IWANT.NE.IBSCN) GO TO 36
0148        KSD=KFND
0149        KSCAN=IWANT
0150     36 CONTINUE
0151        CALL EXEC(24,5,IGCID,0,IERROR)
0152  CCCCCC            PART TWO, INTERPRET THE DATA
0153  CCCCCC LOOK FOR VALLIES AND PEAKS IN CONCENTRATION (TALY)
0154   1000 FORMAT(' EFMP ERR#',I4)
0155     45 LFENET=1+IWANT-IBSCN
0156        NPKS=0
0157  CCCCCC 1ST DET'N WHERE PKS AREA LOCATED. STORE IN CAIN(NPK).
0158        DO 44 J=2,(LFENET-1)
0159        IF((TALY(J).LE.TALY(J-1)).OR.(TALY(J).LT.TALY(J+1)))GOTO44
0160        NPKS=NPKS+1
0161        CAIN(NPKS)=J
0162     44 CONTINUE
0163        IF(NPKS.EQ.0) GO TO 998
0164        SARI=1000.
0165  CCCCCC  >>>>>>> WORK ON EACH PEAK INDIVIDUALLY <<<<<<< CCCCCC
0166        DO 51 I=1,NPKS
0167  CCCCCC FIND BEGINNING OF PK. DEFAULT TO BEG. OF WINDOW.
0168        DO 46 J=1,(CAIN(I)-1)
0169        K=1+CAIN(I)-J
0170        IF((TALY(K-1).GT.TALY(K)).OR.(TALY(K).EQ.0.)) GO TO 47
0171     46 CONTINUE
0172  CCCCCC FIND END OF PK. DEFAULT TO END OF WINDOW.
0173     47 DO 48 L=CAIN(I),LFENET
0174        J=L
0175        IF((TALY(J+1).GT.TALY(J)).OR.(TALY(J).EQ.0.)) GO TO 49
```

```
0176      48 CONTINUE
0177  CCCCCC STORE ABSOLUTE LOC OF PK ST IN CAIN(15+): END IN (30+
0178      49 CAIN(I+15)=K+IBSCN-1
0179         CAIN(I+30)=J+IBSCN-1
0180  CCCCCC INOW INTEGRATE FROM VALLEY TO VALLEY VIA TRAPEZOIDAL METH.
0131         CAIN(I+45)=(TALY(K)+TALY(J))/2.
0182         NN=J-K-1
0183         DO 50 M=1,NN
0184      50 CAIN(I+45)=CAIN(I+45)+TALY(K+M)
0185  CCCCCC DET'N EXACT PEAK LOCATION, STORE IN CAIN(80+NPK)
0186         X=CAIN(I)
0187         X0=(X-1.)*(X-1.)
0188         X1=X*X
0189         X2=(X+1.)*(X+1.)
0190         B=(((X1-X2)*TALY(X-1.))+((X2-X0)*TALY(X))+((X0-X1)*TALY(X+1)
0191        &))/2.
0192         A=(TALY(X+1.)+TALY(X-1)-(2.*TALY(X)))/2.
0193         SCAN=-B/(2.*A)+IBSCN-1
0194         CAIN(60+I)=SCAN
0195  CCCCCC DET'N OBSERVED ARI.  ALSO FIND MIN DELTA ARI.
0196         CAIN(75+I)=CHNGE(SCAN,-1)
0197         IBUF(I)=2
0198         DELRI=ABS(CAIN(75+I)-ARI)
0199         IF(DELRI.GE.SARI) GO TO 51
0200  CCCCCC CLOSEST PK TO ARI BASE REL LOCS IN IBPK & IEPK.
0201         IBPK=K
0202         IEPK=J
0203         SARI=DELRI
0204  CCCCCC MAKE PREPARATIONS FOR PUTTING " >" @ CLOSEST PEAK'S ARI.
0205         IBUF(MOLD)=2
0206         MOLD=I
0207         IBUF(I)=1
0208      51 CONTINUE
0209  CCCCCC REPORT WHAT YOU KNOW.
0210         WRITE(IOUT,101)LIBR,(LIRA(I),I=104,114),ARI,(NOTE(IBUF(I)),
```

```
0211         &CAIN(75+I),CAIN(60+I),CAIN(15+I),CAIN(30+I),CAIN(45+I),
0212         &I=1,NPKS)
0213      101 FORMAT(I4,") ",11A2,F6.1,20(1A2,F6.1,F7.1,I6,I6,I10,/,34X))
0214          GO TO 999
0215      998 WRITE(IOUT,102)LIBR,(LIRA(I),I=104,114),ARI
0216      102 FORMAT(I4,") ",11A2,F6.1,7X,"-",6X,"-",/)
0217      999 CONTINUE
0218          END
0219          END$
**** LIST END ****
$
:LI,S,1,SOPN1


0001   FTN4,L
0002         SUBROUTINE OPEN(IPKN,IFNA,IERROR)
0003   C SUBROUTINE OPEN DEFINES AND OPENS EXTENDED FILE MANAGEMENT
0004   C PACKAGE (EFMP) PACK NUMBER AND FILE NAME.  IF SERIOUS
0005   C ERROR OCCURS, OPEN ABORTS AND REPORTS ERROR IN IERROR.
0006   C IF FILE NOT FOUND, THIS IS REPORTED AND IERROR=-3 RETND
0007   C FOR GCMS, IPKN=111,IFNA=IGCID    FOR LIBR, IPKN=501,IFNA=ARILI
0008         DIMENSION ITRBUF(256),IOPNTB(128),NOTRB(2),IFNA(3)
0009         NOTRB(1)=1
0010         NOTRB(2)=1
0011         CALL EXEC(24,1,IOPNTB,128,ITRBUF,NOTRB,2,IERROR)
0012         IF((IERROR.NE.0).AND.(IERROR.NE.27)) GO TO 2
0013         CALL EXEC(24,4,IFNA,IPKN,1,0,1,IERROR)
0014       2 IF(IERROR.EQ.3) IERROR=-3
0015         IF(IERROR.EQ.25)IERROR=0
0016         IF(IERROR.GT.0)WRITE(1,1)IERROR
0017       1 FORMAT(" OPEN EFMP ERROR $",I3)
0018         RETURN
0019         END
0020         END$
**** LIST END ****
```

```
:LI,S,1,SCHN5

0001   FTN4,L
0002         FUNCTION CHNGE(TI,IR2S)
0003   CHNGE CONVERTS SCAN TO RRI, OR RRI TO SCAN, DEPENDING ON R2S
0004   C IF R2S NEG.,SCN>>>RRI  IF R2S POS.,RRI>>>SCN
0005   C IF TI>LAST STD (OR R2S OMITTED) -2 RETND.
0006   C IF TI<1ST STD: -1 RETND.  CHNGE REQUIRES COMMON CAT(10) &
0007   C MAT(10). SCN TIMES IN CAT, RRI IN MAT. -1 TERMINATES CAT LIST.
0008   C CAT(N)'S SCN TIME MUST CORRESPOND TO MAT(N)'S RRI.
0009         COMMON CAT(10),MAT(10),III(128),MMM(100),CCC(100)
0010         CHNGE=-1
0011         IF(IR2S)5,8,6
0012       5 IF(TI.LT.CAT(1)) RETURN
0013         GO TO 3
0014       6 IF(TI.LT.MAT(1)) RETURN
0015       3 DO 1 I=2,8
0016         IF(CAT(I).EQ.-1) GO TO 8
0017         IF(IR2S)7,9
0018       7 IF(TI.LE.CAT(I)) GO TO 2
0019         GO TO 1
0020       9 IF(TI.LE.MAT(I)) GO TO 4
0021       1 CONTINUE
0022       8 CHNGE=-2
0023         RETURN
0024       2 CHNGE=(((TI-CAT(I-1))/(CAT(I)-CAT(I-1)))*100.)+MAT(I-1)
0025         RETURN
0026       4 CHNGE=(((TI-MAT(I-1))/100.)*(CAT(I)-CAT(I-1)))+CAT(I-1)
0027         RETURN
0028         END
0029         ENDS
**** LIST END ****
```

:LI,S,1,SGETS

```
FTN4,L
      SUBROUTINE GET(IFNA,KSD,KSCAN,KFND,IWANT,LASTM,IERROR)
      COMMON XXX(10),III(10),IBUF(128),MAIN(100),CAIN(100)
      DIMENSION IFNA(3)
      ITMA=1
      IERROR=-1
      IF(KSCAN.GT.IWANT) RETURN
      ISCAN=KSCAN-1
      K=KSD-1
      DO 10 I=1,LASTM
   10 CAIN(I)=0.
CCCCCCCCCCC  GET GCMS SUBROUTINE CCCCCCCCCCCCCCCCCCCCCC
C IONCOUNTS FOR MASSES IN MAIN(I) ARE STORED IN CAIN(I).
C   IFNA  THE FILE NAME OF GCMS RUN
C   KSD   A KNOWN FILE SECTOR, WITH...
C   ISCAN A KNOWN SCAN NUMBER BEGGINNING WITH SECTOR KSD
C   IWANT  THE SCAN NUMBER DESIRED FOR IONCOUNT LOADING
C          IWANT MUST BE GREATER OR EQUAL TO ISCAN
C   LASTM  NUMBER OF MASSES TO FIND IONCOUNTS FOR.
CCCCCCCCCCCCCCC THESE VARS ARE RETURED W/AWNSERS CCCCCCCCCCC
C   KFND   THE ITH RECORD THAT SCAN=IWANT BEGINS ON.
C   IERROR IF ZERO, EVERYTHINGS O.K.
C          IF NEGATIVE, PROBLEM WITH GET SUBPROGRAM
C              -1) IWANT<ISCAN
C              -2) KSD IS NOT THE FIRST SECTOR OF A SCAN,
C                  AND THEREFOR MAY NOT BE A VALID ADDRESS
C          IF POSITIVE, EFMP ERROR (IF 21, END OF GCMS DATA)
C   MASSES IN MAIN() MUST BE ORDERED F/LOW TO HIGHEST.
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
```

```
0030      17 K=K+1
0031         KA=-1
0032         CALL EXEC(24,6,IFNA,K,IBUF,IERROR)
0033         IF(IERROR.NE.0)RETURN
0034         IF((IBUF(4).NE.0).AND.(ISCAN.LT.IWANT)) GO TO 17
0035         IF(((KSCAN-1).EQ.ISCAN).AND.(K.NE.KSD)) GO TO 13
0036   C THIS CHECKS THAT KSD IS FIRST SECTOR OF KSCAN.
0037         IF(IBUF(4).NE.0) GO TO 19
0038         ISCAN=ISCAN+1
0039         IF(ISCAN.LT.IWANT) GO TO 17
0040         IF(ISCAN.GT.IWANT) RETURN
0041         KFND=K
0042         AFCTR=2**IBUF(3)
0043         KA=5
0044      18 KA=KA-2
0045      19 KA=KA+2
0046         IF(KA.GT.128) GO TO 17
0047         IF(IBUF(KA).EQ.-1) RETURN
0048         IF(MAIN(ITMA)-IBUF(KA))14,11,19
0049      11 CAIN(ITMA)=IBUF(KA+1)
0050         IF(CAIN(ITMA).LT.0)CAIN(ITMA)=CAIN(ITMA)+65535.
0051         CAIN(ITMA)=CAIN(ITMA)*AFCTR
0052      14 ITMA=ITMA+1
0053         IF(ITMA.LE.LASTM) GO TO 18
0054         RETURN
0055      13 IERROR=-2
0056         RETURN
0057    C COMPILED VERSION SAVED AS GET5, APPX 650 WDS LONG.
0058         END
0059         END$
**** LIST END ****
```

```
0001  FTN4
0002        PROGRAM TWO
0003        DIMENSION IOPNTB(128),ITRBUF(128),RA(128),MON(128),LIRRI(3)
0004        EXTERNAL OPEN,ERASE,TPLOT,GRAPH
0005  CC COMPILED VERS. OF SUBROUTINES CALLED OPEN1,PLOT,GRAPH.
0006        DATA LIRRI(1)/2HAR/,LIRRI(2)/2HIL/,LIRRI(3)/2HI /,
0007       &IOUT/1/,IN/6/
0008        GO TO 11
0009      2 CALL OPEN(501,LIRRI,IERROR)
0010        IF(IERROR.NE.0)STOP 13
0011        CALL ERASE(1)
0012        WRITE(IOUT,133)
0013    133 FORMAT(" WHAT RANGE OF FILE NUMBERS DO YOU WISH LISTED",
0014       &" IN THE DIRECTORY",/,"INPUT 1ST, 2ND. IF 1ST=-1, ALL.")
0015        READ(IN,*)L1,L2
0016        IF((L1.GT.0).AND.(L2.GT.L1)) GO TO 183
0017        L1=1
0018        L2=1200
0019    183 CONTINUE
0020        WRITE(IOUT,14)
0021     14 FORMAT("FILE#   ARI  $M-RA  COMPOUND NAME")
0022        DO 10 I=L1,L2
0023        CALL EXEC(24,6,LIRRI,I,MON,IERROR)
0024        IF(IERROR.EQ.21)GO TO 10
0025        IF(IERROR.NE.0)WRITE(IOUT,1000)IERROR
0026        ARI=MON(102)/10.
0027        WRITE(IOUT,12)I,ARI,MON(101),(MON(L),L=104,123)
0028     12 FORMAT(I4,")",F6.1,2X,I5,2X,20A2)
0029        IF(MON(102).EQ.-10) GO TO 11
0030     10 CONTINUE
0031     11 WRITE(IOUT,15)
0032        CALL OPEN(501,LIRRI,IERROR)
0033        IF(IERROR.NE.0)WRITE(IOUT,1000)IERROR
0034     15 FORMAT(" INPUT FUNCTION DESIRED;",/,"1) LOOK AT A SPECTRUM"
0035       &,/,"2) DELETE A SPECTRUM",/,"3) MODIFY A LIBRARY SPECTRUM",
```

```
0036        &,/,"4) MANUALLY INPUT A SPECTRUM",/,"5) LIST DIRECTORY",/,
0037        &"6) STOP & QUIT?")
0038         READ(IN,*) ICHOI
0039         IF(ICHOI-2)16,17,303
0040   CCCCCCCCCCCCCC MANUAL INPUT SECTION CCCCCCCCCCCCCZZZZZ
0041     303 IF(ICHOI-4)18,304,305
0042     305 IF(ICHOI-6)2,594
0043     594 STOP 67
0044     304 WRITE(IOUT,306)
0045     306 FORMAT("INPUT COMPOUND NAME")
0046         READ(IN,31)(MON(L),L=104,183)
0047         WRITE(IOUT,310)
0048     310 FORMAT("INPUT ARI")
0049         READ(IN,*)ARI
0050         MON(102)=ARI*10
0051         WRITE(IOUT,311)
0052     311 FORMAT(" HOW MANY MASS-REL.ABUN POINTS DO YOU HAVE?")
0053         READ(IN,*)MON(101)
0054         DO 312 I=1,MON(101)
0055         GO TO 308
0056     307 WRITE(IOUT,309)
0057     309 FORMAT(" ILLEGAL ENTRY, CHECK ORDER AND NUMBER ")
0058     308 WRITE(IOUT,313)I
0059     313 FORMAT("INPUT MASS,RA FOR",I3,"TH PAIR")
0060         READ(IN,*)ITRBUF(I),RA(I)
0061         IF((ITRBUF(I).LT.51).OR.(ITRBUF(I).GT.517).OR.(RA.LT.1).
0062        &OR.(RA.GT.100)) GO TO 307
0063         IF((I.NE.1).AND.(ITRBUF(I).LE.ITRBUF(I-1))) GO TO 307
0064         INTEN=RA(I)
0065         MON(I)=((320-ITRBUF(I))*100)+(INTEN-1)
0066         IF(ITRBUF(I).GT.320)MON(I)=MON(I)-(2*(INTEN-1))
0067     312 CONTINUE
0068         CALL OPEN(501,LIRRI,IERROR)
0069         IF(IERROR.NE.0)PAUSE
0070   CC LOOK FOR APPROPRIATE SPOT TO PUT THE NEW ENTRY.
```

```
0071     123 CONTINUE
0072         DO 500 J=1,2000
0073         CALL EXEC(24,6,LIRRI,J,IOPNTB,IERROR)
0074   CC READ JTH, IF EOF OR RRR>NEW COMPOUNDS FALL THROUGH
0075         IF((IOPNTB(102).NE.-10).AND.(IOPNTB(102).LE.MON(102)))
0076        &GO TO 500
0077         CALL EXEC(24,8,LIRRI,J,MON,IERROR)
0078         IF(IERROR.NE.0)WRITE(IOUT,1000)IERROR
0079   CC PUT LAST WHERE IBUF WAS, THEN PUT IBUF IN LAST.
0080         IF(MON(102).EQ.-10) GO TO 501
0081         DO 503 I=1,128
0082     503 MON(I)=IOPNTB(I)
0083     500 CONTINUE
0084     501 CALL EXEC(24,5,LIRRI,0,IERROR)
0085         GO TO 11
0086   CCCCCCCCCCCCCCC PUT UP A SPECTRUM CCCCCCCCCCCCCCCC
0087      16 WRITE(IOUT,19)
0088      19 FORMAT('INPUT FILE# OF SPECTRA YOU WANT TO SEE, INPUT-1',
0089        &' IF NOT SURE')
0090         READ(IN,*)IFILN
0091         IF(IFILN)2,666,20
0092      20 CALL EXEC(24,6,LIRRI,IFILN,MON,IERROR)
0093         IF(IERROR.NE.0)WRITE(IOUT,1000)IERROR
0094         ARI=MON(102)/10.
0095         CALL ERASE(1)
0096         WRITE(IOUT,2121)LIRRI,IFILN,ARI,(MON(L),L=104,123)
0097    2121 FORMAT('>>>>>>>> LIBRARY NAME:',3A2,'  FILE#',I4,/,
0098        &17X,' RRI',F6.1,' <*>',20A2,'<*>',8/,14X,'100',11X,'150',11X
0099        &'200',11X,'250')
0100         NUOB=MON(101)
0101         NURO=NUOB
0102   CC DECODE MON( ) FOR PLOT, RESTORE SELECTED M-RA INTO ITRB&RA.
0103         DO 69 I=1,NURO
0104         ITRBUF(I)=320-(MON(I)/100)
0105      69 RA(I)=1+IABS(MON(I))-IABS((MON(I)/100)*100)
```

```
0106          I=1
0107          IF(ITRBUF(NUOB).GT.299)I=2
0108          NUOB=NURO
0109          IF(I.EQ.2)WRITE(IOUT,3207)
0110    3207  FORMAT(6/,14X,"350",11X,"400",11X,"450",11X,"500")
0111          WRITE(IOUT,2222)
0112          WRITE(IOUT,2223)(ITRBUF(L),RA(L),L=1,NUOB)
0113    2222  FORMAT(/,7("MASS  RA  "))
0114    2223  FORMAT(7(I4,I4,2X))
0115      70  CALL GRAPH(I)
0116          DO 71 I=1,NUOB
0117          MF=50
0118          IYLO=600
0119          IF(ITRBUF(I).LT.299)GO TO 72
0120          MF=300
0121          IYLO=450
0122      72  IXLO=((ITRBUF(I)-MF)*4)+12
0123          IYHT=RA(I)+IYLO
0124          CALL TPLOT(1,0,IXLO,IYLO)
0125      71  CALL TPLOT(1,1,IXLO,IYHT)
0126          READ(IN,*)DUMMY
0127          GO TO 11
0128  CCCCCCCCCCCC DELETE SECTION CCCCCCCCCCCC
0129      17  WRITE(IOUT,21)
0130      21  FORMAT(" WHAT FILE# IS TO BE DELETED? INPUT -1 TO STOP")
0131          READ(IN,*) IDST
0132          IF(IDST.EQ.-1) STOP 47
0133          NDST=IDST+1
0134          DO 22 INSTR=NDST,1200
0135          INSTP=INSTR-1
0136  C NOW PICK UP SECTOR INSTR AND PUT IT DOUN ON INSTP.
0137          CALL EXEC(24,6,LIRRI,INSTR,MON,IERROR)
0138          IF(IERROR.NE.0)WRITE(IOUT,1000)IERROR
0139          CALL EXEC(24,8,LIRRI,INSTP,MON,IERROR)
0140          IF((MON(102).EQ.-10).OR.(IERROR.EQ.21)) GO TO 23
```

```
0141      22 CONTINUE
0142      23 CALL EXEC(24,5,LIRRI,0,IERROR)
0143         IF(IERROR.NE.0)WRITE(IOUT,1000)IERROR
0144         GO TO 11
0145     666 STOP 45
0146  CCCCCCCCCCCC MODIFY OR STOP SECTION CCCCCCCCCCCCC
0147      18 WRITE(IOUT,24)
0148      24 FORMAT(" WHAT FILE# DO YOU WISH TO MODIFY?, INPUT -1",
0149        &" TO STOP")
0150         READ(IN,*) IFILN
0151         IF(IFILN.LT.1) GO TO 2
0152         CALL OPEN(501,LIRRI,IERROR)
0153         IF(IERROR.NE.0)WRITE(IOUT,1000)IERROR
0154         CALL ERASE(1)
0155         CALL EXEC(24,6,LIRRI,IFILN,MON,IERROR)
0156         IF(IERROR.NE.0)WRITE(IOUT,1000)IERROR
0157         IF(IERROR.NE.0) GO TO 2
0158         WRITE(IOUT,25)
0159      25 FORMAT("DO YOU WISH TO;",/,"1) CHANGE ARI",/,
0160        &"2) CHANGE COMPOUND'S NAME",/,"3) CHANGE M/R.A ?")
0161         READ(IN,*)ICHOI
0162         IF(ICHOI-2)26,27,28
0163      26 ARI=MON(102)/10.
0164         WRITE(IOUT,29)ARI
0165      29 FORMAT(" RRI=",F5.1," INPUT NEW RRI")
0166         READ(IN,*)ARI
0167         MON(102)=ARI*10.
0168         GO TO 312
0169      27 WRITE(IOUT,30)(MON(L),L=104,123)
0170      30 FORMAT(" NAME=",20A2," INPUT NEW NAME")
0171         READ(IN,31)(MON(L),L=104,123)
0172      31 FORMAT(20A2)
0173  CC MAKE SURE NEW ENTRY'S ARI IS PROPERLY PLACED.
0174         WRITE(IOUT,494)
0175     494 FORMAT(" DESTROY OLD NAME'S FILE? 1=YES, 2=NO")
```

```
0176          READ(IN,*)IDES
0177          IF(IDES.NE.1) GO TO 123
0178          CALL EXEC(24,8,LIRRI,IFILN,MON,IERROR)
0179          IF(IERROR.NE.0)WRITE(IOUT,1000)IERROR
0180          CALL EXEC(24,5,LIRRI,0,IERROR)
0181          GO TO 11
0182    CCCCCCCCCCCCC CHAGE MASS-REL.ABUN PAIR SECTION CCCCCCCCCCCCCCC
0183       28 CONTINUE
0184          CALL OPEN(501,LIRRI,IERROR)
0185          IF(IERROR.NE.0)WRITE(IOUT,1000)IERROR
0186          CALL ERASE(1)
0187          CALL EXEC(24,6,LIRRI,IFILN,MON,IERROR)
0188          IF(IERROR.NE.0)WRITE(IOUT,1000)IERROR
0189          DO 169 I=1,MON(101)
0190          ITRBUF(I)=320-(MON(I)/100)
0191      169 RA(I)=1+IABS(MON(I))-IABS((MON(I)/100)*100)
0192          WRITE(IOUT,2222)
0193          WRITE(IOUT,2223)(ITRBUF(L),RA(L),L=1,MON(101))
0194          WRITE(IOUT,34)
0195       34 FORMAT(' DO YOU WANT TO 1) ADD',/,16X,'2) DELETE A M-RA'
0196         &,' PAIR',/,16X,'3) STOP EDITING?')
0197          READ(IN,*)ICHOI
0198          IF(ICHOI-2) 35,36,37
0199       37 GO TO 11
0200       35 WRITE(IOUT,38)
0201       38 FORMAT('INPUT MASS, REL-ABUN TO ADD INTO LIBRARY')
0202          READ(IN,*)MAS,REL
0203          IF((MAS.LT.51).OR.(MAS.GT.517).OR.(REL.LT.1).OR.(REL.GT.100)
0204         &) GO TO 34
0205          DO 40 I=1,MON(101)
0206          IF(MAS.EQ.ITRBUF(I)) GO TO 28
0207          IF(ITRBUF(I).GT.MAS) GO TO 41
0208       40 CONTINUE
0209       41 CONTINUE
0210          MON(101)=MON(101)+1
```

```
      DO 42 J=I,MON(101)
      KMA=ITRBUF(J)
      KRA=RA(J)
      ITRBUF(J)=MAS
      RA(J)=REL
      MAS=KMA
   42 REL=KRA
      DO 168 I=1,MON(101)
      MON(I)=((320-ITRBUF(I))*100)+(RA(I)-1)
      IF(ITRBUF(I).GT.320) MON(I)=MON(I)-(2*(RA(I)-1))
  168 CONTINUE
      GO TO 676
   36 WRITE(IOUT,101)
  101 FORMAT("WHAT IS THE MASS OF THE PAIR YOU WISH TO DELETE?")
      READ(IN,*)MAS
      DO 102 I=1,MON(101)
      IF(ITRBUF(I).EQ.MAS) GO TO 103
  102 CONTINUE
      GO TO 28
  103 DO 104 J=I,MON(101)
  104 MON(J)=MON(J+1)
      MON(101)=MON(101)-1
  676 CONTINUE
      CALL EXEC(24,8,LIRRI,IFILN,MON,IERROR)
      IF(IERROR.NE.0)WRITE(IOUT,1000)IERROR
      CALL EXEC(24,5,LIRRI,0,IERROR)
      IF(IERROR.NE.0)WRITE(IOUT,1000)IERROR
      GO TO 28
   33 CALL EXEC(24,8,LIRRI,IFILN,MON,IERROR)
      IF(IERROR.NE.0)WRITE(IOUT,1000)IERROR
      CALL EXEC(24,5,LIRRI,0,IERROR)
      IF(IERROR.NE.0)WRITE(IOUT,1000)IERROR
      GO TO 18
 1000 FORMAT(" EFMP ERROR #",I4)
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
```

```
0246  CCCCCC PROGRAM TWO REQUIRES OPEN1, PLOT AND GRAPH
0247  CCCCCC SUBROUTINES TO BE  ATTACHED WHEN THIS
0248  CCCCCC PROGRAM IS LOADED.
0249  CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
0250        END
0251        END$
**** LIST END ****
:
:LI,S,1,SGRAPH


0001  FTN4,L
0002        SUBROUTINE GRAPH(NPLOTS)
0003        IFIF=4
0004        ITEN=1
0005        IFIV=4
0006        DO 400 MA=12,1004,4
0007        IFIV=IFIV+1
0008        MAN=0
0009        IF(IFIV.NE.5) GO TO 401
0010        MAN=MAN-5
0011        IFIV=0
0012        ITEN=ITEN+1
0013        IF(ITEN.NE.2) GO TO 401
0014        ITEN=0
0015        MAN=MAN-5
0016        IFIF=IFIF+1
0017        IF(IFIF.NE.5) GO TO 401
0018        IFIF=0
0019        MAN=MAN-5
0020    401 CONTINUE
0021        CALL TPLOT(1,0,MA,601)
0022        CALL TPLOT(1,1,MA,(600+MAN))
0023        IF(NPLOTS.NE.2) GO TO 400
0024        CALL TPLOT(1,0,MA,449)
```

```
0025          CALL TPLOT(1,1,MA,(450+MAN))
0026      400 CONTINUE
0027          CALL TPLOT(1,0,7,700)
0028          CALL TPLOT(1,1,7,600)
0029          CALL TPLOT(1,1,1015,600)
0030          CALL TPLOT(1,1,1015,700)
0031          IF(NPLOTS.NE.2) RETURN
0032          CALL TPLOT(1,0,7,550)
0033          CALL TPLOT(1,1,7,450)
0034          CALL TPLOT(1,1,1015,450)
0035          CALL TPLOT(1,1,1015,550)
0036          RETURN
0037          END
0038          END$
**** LIST END ****
@
:LI,S,1,STHR


0001    FTN4
0002          PROGRAM THREE
0003          COMMON ITRBUF(256),IOPNTB(128)
0004          DIMENSION RA(256),CAIN(50,5),MON(256),ABUN(85),IBUF(128)
0005         &,LAST(128),AMASS(40),MAIN(50),IGCID(3),LIRRI(3)
0006          EQUIVALENCE (LAST,AMASS)
0007          EQUIVALENCE (RA,CAIN)
0008          EQUIVALENCE (MON,ABUN),(MON(171),IBUF)
0009          DATA IOUT/1/,IN/6/,KSD/5/,KSCAN/1/,LUN/2/,ITRK/190/
0010         &,ISECT/0/,LIRRI(1)/2HAR/,LIRRI(2)/2HIL/,LIRRI(3)/2HI /
0011    CC LUN=2 MEANS LOWER DISK IS USED AS SCRATCH AREA & GCMS IS TOP.
0012        2 WRITE(IOUT,1001)
0013          READ(IN,1002)IGCID
0014          CALL OPEN(111,IGCID,IERROR)
0015          IF(IERROR)2,1,3
0016        3 PAUSE 1
```

```
0017      1 WRITE(IOUT,700)
0018    700 FORMAT(' INPUT SCAN NUMBER CLOSEST TO COMPOUND MAX.')
0019        READ(IN,*)ICENT
0020        IF((ICENT.LT.2).OR.(ICENT.GT.1570)) GO TO 1
0021        IBSCN=ICENT-2
0022        IESCN=ICENT+2
0023      4 WRITE(IOUT,702)
0024    702 FORMAT('INPUT DESIRED MASS RANGE: LO,HI (LIMITS:51,517)')
0025        READ(IN,*)M1,M2
0026        IF((M2.LE.M1).OR.(M1.LT.51).OR.(M2.GT.517))GO TO 4
0027  CC DO 100 MASSES AT A TIME FOR EACH OF THE SCANS OBSERVED.
0028        DO 300 MA=M1,M2,51
0029        NUOB=0
0030        MB=MA+50
0031        IF(MB.GT.M2)MB=M2
0032  CC LOAD DESIRED MASSES INTO MAIN(), COUNT HOW MANY U/LASTM.
0033        LASTM=0
0034        DO 12 I=MA,MB
0035        LASTM=LASTM+1
0036     12 MAIN(LASTM)=I
0037  CC CLEAR OUT ION INTENSITY BUFFER CAIN VIA RA.
0038        DO 13 I=1,256
0039     13 RA(I)=0.
0040        IRS=0
0041  CC NOW GO LOOKING FOR INTENSITIES FOR SCANS NEAR CENTER.
0042        DO 10 IWANT=IBSCN,IESCN
0043        IRS=IRS+1
0044  CC THIS IS GET SUBROUTINE
0045        ITMA=1
0046        K=KSD-1
0047        ISCAN=KSCAN-1
0048     17 K=K+1
0049        KA=-1
0050        CALL EXEC(24,6,IGCID,K,IBUF,IERROR)
0051        IF(IERROR.EQ.0) GO TO 16
```

```
      WRITE(IOUT,1000)IERROR
      PAUSE 2
   16 IF((IBUF(4).NE.0).AND.(ISCAN.LT.IWANT)) GO TO 17
      IF(((KSCAN-1).EQ.ISCAN).AND.(K.NE.KSD)) PAUSE 3
CC PAUSE 3 INDICATES KSTOR IS NOT A BEGINNING SCAN ADDRESS.
      IF(IBUF(4).NE.0) GO TO 19
      ISCAN=ISCAN+1
      IF(ISCAN.LT.IWANT) GO TO 17
      IF(ISCAN.GT.IWANT) PAUSE 4
CC PAUSE 4 INDICATES  ERROR IN GET ROUTINE
      KFND=K
      AFCTR=2**IBUF(3)
      KA=5
   18 KA=KA-2
   19 KA=KA+2
      IF(KA.GT.128) GO TO 17
      IF(IBUF(KA).EQ.-1) GO TO 5
      IF(MAIN(ITMA)-IBUF(KA))14,11,19
   11 CAIN(ITMA,IRS)=IBUF(KA+1)
      IF(IBUF(KA+1).LT.0)CAIN(ITMA,IRS)=CAIN(ITMA,IRS)+65535.
      CAIN(ITMA,IRS)=CAIN(ITMA,IRS)*AFCTR
   14 ITMA=ITMA+1
      IF(ITMA.LE.LASTM) GO TO 18
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
    5 IF(IRS.NE.1) GO TO 10
      KSD=KFND
      KSCAN=IWANT
   10 CONTINUE
      DO 15 KK=1,85
   15 IBUF(KK)=0
      DO 202 I=1,LASTM
      DO 100 J=2,4
      Y0=CAIN(I,(J-1))
      Y1=CAIN(I,J)
      Y2=CAIN(I,(J+1))
```

```
0037          IF((Y1.LE.Y0).OR.(Y1.LT.Y2)) GO TO 100
0088          IF((Y0.EQ.0).AND.(Y2.EQ.0)) GO TO 100
0089    CC IF DOWN HERE, MAX AT J.
0090          X0=J-1
0091          X1=J
0092          X2=J+1
0093          B=((((X1*X1)-(X2*X2))*Y0)+(((X2*X2)-(X0*X0))*Y1)
0094         &+(((X0*X0)-(X1*X1))*Y2))/2.
0095          A=(Y2+Y0-(2.*Y1))/2.
0096          XMP=-B/(2.*A)
0097          YMP=((((XMP-X1)*(XMP-X2))/2.)*Y0)+(((XMP-X0)*(XMP-X2))
0098         &*(-Y1))+((((XMP-X0)*(XMP-X1))/2.)*Y2)
0099          IXMP=(XMP*10)-14
0100          IF((IXMP.GT.0).AND.(IXMP.LT.37)) GO TO 201
0101          WRITE(IOUT,666)(MAIN(N),(CAIN(LN,M),M=1,5),N=1,LASTM)
0102      666 FORMAT(' M/E=',I4,4X,'RSN(1-5):',5F9.0)
0103          WRITE(IOUT,666)MAIN(I)
0104          PAUSE 6
0105    CC PAUSE 6 MEANS PK. OBSERVED OUT OF RANGE.
0106          GO TO 100
0107    CC IXMP REPS. FRACTION BETWEEN IBSCN-ICENT-IESCN.
0108      201 NUOB=NUOB+1
0109          IF(NUOB.GT.85) PAUSE 7
0110          IBUF(NUOB)=MAIN(I)+(IXMP*1000)
0111          AMASS(IXMP)=AMASS(IXMP)+YMP
0112          ABUN(NUOB)=CAIN(I,(J-1))+CAIN(I,J)+CAIN(I,(J+1))
0113      100 CONTINUE
0114      202 CONTINUE
0115          ISECT=ISECT+2
0116      300 CALL BWRIT(MON,256,LUN,ITRK,ISECT,0)
0117    CC NOW PRESENT AMASS DATA, ASK OPERATOR WHAT MC PK TIMES TO INCL.
0118      888 NUOB=0
0119          CALL ERASE(1)
0120    CC GENERATE REAL T.U.S
0181          DO 50 I=1,36
```

```
0122      50 ABUN(I)=(I/10.)+IBSCN
0123         WRITE(IOUT,52)ABUN(IBEG),ABUN(IEND)
0124      52 FORMAT("FORMAR MC PK RANGE",2F7.1,/,4("TU  SCAN IONSUN X "))
0125         WRITE(IOUT,54)(L,ABUN(L),AMASS(L),L=1,36)
0126      54 FORMAT(4(I2,F6.1,I7," X "))
0127         WRITE(IOUT,56)
0128      56 FORMAT(" INPUT DESIRED INCLUSION RANGE FOR MC PKS",/,
0129        &" INPUT TU#1,TU#2.  IF TU#1=-1, PROG HALTS.")
0130         READ(IN,*)IBEG,IEND
0131         IF(IBEG.EQ.-1) STOP 7
0132         IF((IEND.LT.IBEG).OR.(IEND.GT.36).OR.(IBEG.LT.1))GO TO 888
0133         WRITE(IOUT,58)
0134      58 FORMAT("WHEN SPECTRA IS PRESENTED,",/,"RETN 1 FOR M/E",
0135        &," & R.A. TABLE NEXT TIME,",/,"     0 FOR NO M/E TABLE",
0136        &,/,"      7 TRANSFERS SPECTRUM DATA TO DISK FOR LIBRARY USE"
0137        &,/," NOW INPUT LARGEST DESIRED MASS FOR SPECTRUM")
0138         READ(IN,*)LAMA
0139         IF(LAMA.GT.M2)LAMA=M2
0140         WRITE(IOUT,59)MIN
0141      59 FORMAT(" MINIMUM R.A.=",I4," INPUT NEW FILTER.")
0142         READ(IN,*)MIN
0143         IF((MIN.LT.1).OR.(MIN.GT.100))MIN=1
0144         CALL ERASE(1)
0145         WRITE(IOUT,57) IQCID,ABUN(IBEG),ABUN(IEND),MIN,M1,LAMA
0146      57 FORMAT(5X,"(*) PROGRAM PARE: MS CLEANUP VIA MC PROFILE",
0147        &" ANALYSES (*)",/,"QC ID:",3A2,29X,"SCAN RANGE:",2F8.1,
0148        &/,"MIN RA.",I3,"%",30X,"MASS RANGE:",2I8,7/,14X,"100",11X,
0149        &"150",11X,"200",11X,"250")
0150         DO 61 I=1,256
0151         ITRBUF(I)=0
0152      61 RA(I)=0.
0153         I100=IBEG*1000
0154         I200=(IEND*1000)+518
0155         BIG=0.
0156         DO 60 K=2,ISECT,2
```

```
0157          CALL BREAD(MON,256,LUN,ITRK,K,0)
0158    CC LOAD AAIN( )&MAIN( ) VIA TIM WITH TIMES,INTENS&MASSES.
0159          DO 64 I=1,85
0160          IF((IBUF(I).LT.I100).OR.(IBUF(I).GT.I200))GO TO 64
0161    CC IF IT COMES DOWN THIS WAY, MASS FOUND IN TIME RANGE.
0162          IF(NUOB.LT.257) GO TO 62
0163          PAUSE 11
0164          GO TO 888
0165    CC PAUSE 11 MEANS TOO MANY M/E PEAKS (>256).
0166       62 NUOB=NUOB+1
0167          ITRBUF(NUOB)=IBUF(I)-((IBUF(I)/1000)*1000)
0168          IF(ITRBUF(NUOB).EQ.ITRBUF(NUOB-1))PAUSE 10
0169    CC PAUSE 10 MEANS WIDE ENOUGH INCLUSION RANGE TO SEE MASS PK 2X.
0170          IF(ITRBUF(NUOB).GT.LAMA) GO TO 65
0171          RA(NUOB)=ABUN(I)
0172          IF(RA(NUOB).GT.BIG)BIG=RA(NUOB)
0173       64 CONTINUE
0174       60 CONTINUE
0175       65 IF(NUOB.EQ.0) GO TO 888
0176          NURO=0
0177          X1=100./BIG
0178    CC CODE MASS(ITRBUF( )) & RA(RA( )) INTO 1 INTERGER WORD (MON( ))
0179          DO 68 I=1,NUOB
0180          INTEN=RA(I)*X1
0181          IF(INTEN.LT.MIN) GO TO 68
0182          NURO=NURO+1
0183          MON(NURO)=((320-ITRBUF(I))*100)+(INTEN-1)
0184          IF(ITRBUF(I).GT.320)MON(NURO)=MON(NURO)-(2*(INTEN-1))
0185       68 CONTINUE
0186    CC DECODE MON( ) FOR PLOT, RESTORE SELECTED M-RA INTO ITRB&RA.
0187          DO 69 I=1,NURO
0188          ITRBUF(I)=320-(MON(I)/100)
0189       69 RA(I)=1+IABS(MON(I))-IABS((MON(I)/100)*100)
0190          I=1
0191          IF(ITRBUF(NUOB).GT.299)I=2
```

```
0192          NUOB=NURO
0193          IF(I.EQ.2)WRITE(IOUT,3207)
0194    3207  FORMAT(6/,14X,"350",11X,"400",11X,"450",11X,"500")
0195          IF(NOY.NE.1) GO TO 70
0196          WRITE(IOUT,2222)
0197          WRITE(IOUT,2223)(ITRBUF(L),RA(L),L=1,NUOB)
0198    2222  FORMAT(/,7("MASS  RA  "))
0199    2223  FORMAT(7(I4,I4,2X))
0200      70  CALL GRAPH(I)
0201          DO 71 I=1,NUOB
0202          MF=50
0203          IYLO=600
0204          IF(ITRBUF(I).LT.299)GO TO 72
0205          MF=300
0206          IYLO=450
0207      72  IXLO=((ITRBUF(I)-MF)*4)+12
0208          IYHT=RA(I)+IYLO
0209          CALL TPLOT(1,0,IXLO,IYLO)
0210      71  CALL TPLOT(1,1,IXLO,IYHT)
0211          READ(IN,*) NOY
0212          IF((NOY.NE.7).OR.(NUOB.GT.100))GO TO 888
0213          CALL ERASE(1)
0214          MON(101)=NUOB
0215      80  WRITE(IOUT,82)
0216      82  FORMAT("INPUT RRI, THEN COMPOUND NAME:")
0217          READ(IN,*)X1
0218          MON(102)=X1*10
0219          READ(IN,84)(MON(L),L=104,123)
0220      84  FORMAT(20A2)
0221    CC PREPARE TO ENTER NEW COMPOUND IN LIBRARY.
0222          DO 88 I=1,128
0223      88  LAST(I)=MON(I)
0224          CALL OPEN(501,LIRRI,IERROR)
0225          IF(IERROR.NE.0)PAUSE
0226    CC LOOK FOR APPROPRIATE SPOT TO PUT THE NEW ENTRY.
```

```
0227          DO 500 J=1,2000
0228          CALL EXEC(24,6,LIRRI,J,IBUF,IERROR)
0229    CC READ JTH, IF EOF OR RRR>NEW COMPOUNDS FALL THROUGH
0230          IF((IBUF(102).NE.-10).AND.(IBUF(102).LE.LAST(102)))GOTO500
0231          CALL EXEC(24,8,LIRRI,J,LAST,IERROR)
0232          IF(IERROR.NE.0)WRITE(IOUT,1000)IERROR
0233    CC PUT LAST WHERE IBUF WAS, THEN PUT IBUF IN LAST.
0234          IF(LAST(102).EQ.-10)GO TO 501
0235          DO 503 I=1,128
0236      503 LAST(I)=IBUF(I)
0237      500 CONTINUE
0238      501 CALL EXEC(24,5,LIRRI,0,IERROR)
0239          STOP 2
0240    CC PROGRAM THR REQUIRES OPEN3, PLOT & GRAPH IN LOADING.
0241     1000 FORMAT("EFMP ERROR NO.",I3)
0242     1001 FORMAT("ENTER AN EXISTING GCID AS FIVE CHARACTERS")
0243     1009 FORMAT(8F8.1)
0244     1002 FORMAT(3A2)
0245          END
0246          END$
**** LIST END ****
```

```
:LI,S,1,SOPN3

0001   FTN4,L
0002         SUBROUTINE OPEN(IPKN,IFNA,IERROR)
0003   C SUBROUTINE OPEN DEFINES AND OPENS EXTENDED FILE MANAGEMENT
0004   C PACKAGE (EFMP) PACK NUMBER AND FILE NAME.  IF SERIOUS
0005   C ERROR OCCURS, OPEN ABORTS AND REPORTS ERROR IN IERROR.
0006   C IF FILE NOT FOUND, THIS IS REPORTED AND IERROR=-3 RETND
0007   C FOR GCMS, IPKN=111,IFNA=IGCID    FOR LIBR, IPKN=501,IFNA=ARILI
0008         COMMON ITRBUF(256),IOPNTB(128)
0009         DIMENSION IFNA(3),NOTRB(2)
0010         NOTRB(1)=1
0011         NOTRB(2)=1
0012         CALL EXEC(24,1,IOPNTB,128,ITRBUF,NOTRB,2,IERROR)
0013         IF((IERROR.NE.0).AND.(IERROR.NE.27)) GO TO 2
0014         CALL EXEC(24,4,IFNA,IPKN,1,0,1,IERROR)
0015       2 IF(IERROR.EQ.3) IERROR=-3
0016         IF(IERROR.EQ.25)IERROR=0
0017         IF(IERROR.GT.0)WRITE(1,1)IERROR
0018       1 FORMAT(' OPEN EFMP ERROR #',I3)
0019         RETURN
0020         END
0021         END$
**** LIST END ****
```